LECTURE NOTES IN PHYSICS

D. Reguera L.L. Bonilla J.M. Rubí (Eds.)

Coherent Structures in Complex Systems

Springer

Lecture Notes in Physics

Editorial Board

R. Beig, Wien, Austria
J. Ehlers, Potsdam, Germany
U. Frisch, Nice, France
K. Hepp, Zürich, Switzerland
W. Hillebrandt, Garching, Germany
D. Imboden, Zürich, Switzerland
R. L. Jaffe, Cambridge, MA, USA
R. Kippenhahn, Göttingen, Germany
R. Lipowsky, Golm, Germany
H. v. Löhneysen, Karlsruhe, Germany
I. Ojima, Kyoto, Japan
H. A. Weidenmüller, Heidelberg, Germany
J. Wess, München, Germany
J. Zittartz, Köln, Germany

Springer
Berlin
Heidelberg
New York
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

Physics and Astronomy ONLINE LIBRARY
http://www.springer.de/phys/

Editorial Policy

The series *Lecture Notes in Physics* (LNP), founded in 1969, reports new developments in physics research and teaching -- quickly, informally but with a high quality. Manuscripts to be considered for publication are topical volumes consisting of a limited number of contributions, carefully edited and closely related to each other. Each contribution should contain at least partly original and previously unpublished material, be written in a clear, pedagogical style and aimed at a broader readership, especially graduate students and nonspecialist researchers wishing to familiarize themselves with the topic concerned. For this reason, traditional proceedings cannot be considered for this series though volumes to appear in this series are often based on material presented at conferences, workshops and schools (in exceptional cases the original papers and/or those not included in the printed book may be added on an accompanying CD ROM, together with the abstracts of posters and other material suitable for publication, e.g. large tables, colour pictures, program codes, etc.).

Acceptance

A project can only be accepted tentatively for publication, by both the editorial board and the publisher, following thorough examination of the material submitted. The book proposal sent to the publisher should consist at least of a preliminary table of contents outlining the structure of the book together with abstracts of all contributions to be included.
Final acceptance is issued by the series editor in charge, in consultation with the publisher, only after receiving the complete manuscript. Final acceptance, possibly requiring minor corrections, usually follows the tentative acceptance unless the final manuscript differs significantly from expectations (project outline). In particular, the series editors are entitled to reject individual contributions if they do not meet the high quality standards of this series. The final manuscript must be camera-ready, and should include both an informative introduction and a sufficiently detailed subject index.

Contractual Aspects

Publication in LNP is free of charge. There is no formal contract, no royalties are paid, and no bulk orders are required, although special discounts are offered in this case. The volume editors receive jointly 30 free copies for their personal use and are entitled, as are the contributing authors, to purchase Springer books at a reduced rate. The publisher secures the copyright for each volume. As a rule, no reprints of individual contributions can be supplied.

Manuscript Submission

The manuscript in its final and approved version must be submitted in camera-ready form. The corresponding electronic source files are also required for the production process, in particular the online version. Technical assistance in compiling the final manuscript can be provided by the publisher's production editor(s), especially with regard to the publisher's own Latex macro package which has been specially designed for this series.

Online Version/ LNP Homepage

LNP homepage (list of available titles, aims and scope, editorial contacts etc.):
http://www.springer.de/phys/books/lnpp/

LNP online (abstracts, full-texts, subscriptions etc.):
http://link.springer.de/series/lnpp/

D. Reguera L.L. Bonilla J.M. Rubí (Eds.)

Coherent Structures in Complex Systems

Selected Papers of the XVII Sitges Conference
on Statistical Mechanics
Held at Sitges, Barcelona, Spain, 5-9 June 2000

Springer

Editors

David Reguera
José Miguel Rubí
Department de Física Fonamental, Universitat de Barcelona
Diagonal 647, Barcelona 08028, Spain

Luis López Bonilla
Departamento de Matemáticas
Escuela Politécnica Superior
Universidad Carlos III de Madrid
Avenida de la Universidad 30, Leganés 28911, Spain

Library of Congress Cataloging-in-Publication Data applied for.

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Coherent structures in complex systems : selected papers of the XVII Sitges Conference on Statistical Mechanics, held at Sitges, Barcelona, Spain, 5 - 9 June 2000 / D. Reguera ... (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 2001
(Lecture notes in physics ; Vol. 567)
(Physics and astronomy online library)
ISBN 3-540-41705-2

ISSN 0075-8450
ISBN 3-540-41705-2 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

http://www.springer.de

© Springer-Verlag Berlin Heidelberg 2001
Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by the authors/editors
Camera-data conversion by Steingraeber Satztechnik GmbH Heidelberg
Cover design: *design & production*, Heidelberg

Printed on acid-free paper
SPIN: 10797201 57/3141/du - 5 4 3 2 1 0

Preface

A rich variety of real-life physical problems which are still poorly understood are of a nonlinear nature. Examples include turbulence, weather forecasting, granular flows, detonations and flame propagation, fracture dynamics, and a wealth of new biological and chemical phenomena which are being discovered. On these problems, there has been significant historical interest, as well as a considerable amount of experimental data and research into industrial applications during the last 20 years. However, the essentials of many underlying physical mechanisms remain unexplained despite the high potential interest from the theoretical and applied points of view and the considerable progress achieved during recent years. The current theoretical understanding of these subjects is based on similar techniques and approaches. This allows for a common language and the existence of a unifying background for the manifold phenomena arising from nonlinearities.

Most interesting among the manifestations of nonlinearity are coherent structures. In turbulence, numerical evidence of high-vorticity small-size filaments in simulations of the full Navier–Stokes equations has given great theoretical and experimental impetus to this field and many new ideas about the role of coherent structures are now the subject of intense scrutiny. In practical applications, such as internal combustion engines, it is of fundamental importance to know how the speed of propagation of a flame front is affected by the turbulence intensity. There are diverse regimes controlled by the ratio of the Kolmogorov scale of the turbulent flow to the flame width. At the relatively well-understood limit of large values of this ratio, the phenomenon of turbulent flame wrinkling appears. In the other limit (of much greater practical importance), flame extinction, saturation of the flame velocity with turbulence intensity and other interesting phenomena may appear. Theoretical understanding of these phenomena is still poor and strong controversy has arisen. The formation of structures is also particularly important in gaseous detonations, where much discussion has followed new theoretical advances in understanding the interplay between fluid dynamics, chemistry and heat release in a detonation wave, in both one-dimensional and multidimensional cases. Pattern formation and spatio-temporal structures are also prominent in fluid flow, dendritic growth, and chemical and biological phenomena. Lastly, granular flow and fracture dynamics are developing fields posing very challenging problems with important nonlinear and statistical aspects which will no doubt be the object of a great deal of attention and work in coming years.

Motivated by these considerations, we gathered several leading experts for the XVII Sitges Conference. This book contains a number of reviews and contributions reporting on the state of the art as it concerns the role of coherent structures and patterns in the above-listed problems of nonlinear science.

The conference was the second of a series of two Euroconferences focusing on the topic *Nonlinear Phenomena in Classical and Quantum Systems*. It was sponsored by the UE (Euroconference) and by institutions which generously provided financial support: DGES of the Spanish Government, CIRIT of the Generalitat of Catalonia, Universitat de Barcelona and Universidad Carlos III de Madrid. The city of Sitges allowed us, as usual, to use the Palau Maricel as the lecture hall.

Finally, we are also very grateful to all those who collaborated in the organization of the event, Prof. C. Dopazo, Drs. A. Pérez-Madrid and T. Alarcón, and M. González, I. Santamaría-Holek, I. Pérez-Castillo, and O. Sánchez.

Barcelona, November 2000 *The Editors*

Contents

Part I Turbulence

Description and Dynamics of Vortical Structures of Turbulence
Shigeo Kida 3

Self-Sustaining Mechanisms of Wall Turbulence
Alexander J. Smits, Carl Delo 17

The Largest Scales in Turbulent Flow: The Structures of the Wall Layer
Javier Jiménez, Mark Simens 39

Patterning and Transition to Turbulence in Subcritical Systems: The Case of Plane Couette Flow
Paul Manneville, Olivier Dauchot 58

Effects of the Upstream Conditions on the Mean Statistics of Turbulent Boundary Layers with Zero Pressure Gradient
Luciano Castillo, Gunnar Johansson, Erik Eisenmen 80

Evolution of the Large-Scale Structures in the 'Far-Field' of Turbulent Shear Flows
Dan Ewing 92

Quantum Kinetic Model of Turbulence
Amador Muriel 103

The Transport of Small Particles by a Fluid
Armando Babiano, Julyan H.E. Cartwright, Oreste Piro, Antonello Provenzale 114

Part II Combustion

Nonlinear Dynamics of Wrinkled Premixed Flames and Related Statistical Problems
Guy Joulin, Gaël Boury, Pierre Cambray, Yves D'Angelo, Karl Joulain . . 127

Experimental Studies of Laminar Flame Instabilities
Geoff Searby, Jean-Marie Truffaut 159

Gaseous Detonations
Paul Clavin 182

The Vortical Structure of Flame Spreading over Liquid Fuels
Francisco J. Higuera, José C. Antoranz, Vladimir Sankovitch, José L. Castillo, Eugenio Degroote, Pedro L. Garcia-Ybarra 190

Part III Patterns in Biology, Chemical Reactions and Fluid Flow

Applications of Mathematical Modelling to Biological Pattern Formation
Philip K. Maini 205

The Artistry of Bacterial Colonies and the Antibiotic Crisis
Ido Golding, Eshel Ben-Jacob 218

Periodicity in Age-Resolved Populations
Sergei Esipov 242

Nonequilibrium Nanostructures in Condensed Reactive Systems
Alexander S. Mikhailov, Michael Hildebrand, Gerhard Ertl 252

Pattern Formation in Electric Discharges
Ute Ebert, Manuel Arrayás 270

Evidence for Eigenfrequencies in Dendritic Growth Dynamics
Jeffrey C. LaCombe, Matthew B. Koss, Cindie Giummarra, Julie E. Frei, Afina O. Lupulescu, Martin E. Glicksman 283

Patterns in the Bulk and at the Interface of Liquid Crystals
Ágnes Buka, Tamás Börzsönyi, Nándor Éber, Tibor Tóth-Katona 298

Inviscid Two-Dimensional Fluid Dynamics Experiments with Magnetized Electron Columns
Joel Fajans, Daniel Durkin 319

Interaction of Nearly-Inviscid, Multi-mode Faraday Waves and Mean Flows
María J. Higuera, José M. Vega, Edgar Knobloch 328

Stationary and Oscillatory Flow of a Charged Liquid Around a Blade Electrode
Francisco J. Higuera 338

Part IV Granular Media and Fractures

Science in the Sandbox: Fluctuations, Friction and Instabilities
Robert P. Behringer, Eric Clément, Junfei Geng, Dan Howell, Ljubinko Kondic, Guy Metcalfe, Corey O'Hern, Guillaume Reydellet, Sarath Tennakoon, Loic Vanel, Christian Veje 351

Cluster Instability in Freely Evolving Granular Gases
J. Javier Brey 392

Granular Gases – The Early Stage
Nikolai V. Brilliantov, Thorsten Pöschel 408

Experimental and Theoretical Study of the Gravity Induced Granular Flow in Two-Dimensional Silos
Abraham Medina, César Treviño, Pedro J. Hernando, Luis L. Bonilla 420

Dynamics of Fracture
Farid F. Abraham 436

Avalanches and Damage Clusters in Fracture Processes
Stefano Zapperi, Purusattam Ray, H. Eugene Stanley, Alessandro Vespignani 452

List of Participants 461

Index 467

Description and Dynamics of Vortical Structures of Turbulence

Shigeo Kida

Theory and Computer Simulation Center,
National Institute for Fusion Science, Toki, 509-5292, Japan

Abstract. The vortical structure of turbulence is analyzed in terms of the low-pressure vortex, the central axis of which is given by loci of pressure minimum and the core is defined as concave regions of pressure. Since the position and direction of vortex axes are specified, we can analyze the vorticity field by decomposing the axial and cross-axial components. It is found that the fundamental vortical structure — the elementary vortex — is a tubular vortex core with strong axial vorticity accompanied with spirally wrapped vortex layer of cross-axial vorticity. These surrounding layers can contribute more than the tubular cores to energy dissipation because of their wider spatial extent.

1 Introduction

Turbulence is a complex motion of fluid both in space and time. In order to understand the dynamics it is indispensable to see what is actually happening. The apparent structure may be different depending upon the physical quantities to visualize with. The isosurfaces of high-vorticity, low-pressure, high-Laplacian of pressure, etc. have often been used for visualization of flow structures. It is in fact by these methods that the common existence of tubular swirling coherent motions of concentrated vorticity were discovered. However, the isosurface representation inevitably suffers from arbitrariness. The resulting structure changes its shape depending on the thresholds to be chosen (though it is possible to get a useful insight by comparing the isosurfaces visualized at different levels). This choice is usually arbitrary and it is hard to find a reasonable criterion of the threshold in general. A visualization method without arbitrariness is desirable.

In view of its universal existence in turbulence we take the tubular swirling vortex as a reference object to describe flows. We propose to characterize the vortex as low-pressure regions rather than high-vorticity. This idea is based upon the fact that the pressure typically takes a minimum at the center of swirling motions and that high-vorticity regions do not necessarily represent swirling motions (the relative magnitude of the vorticity and strain rate determines the swirl condition [14]). The pressure is surveyed in the whole flow field and the loci of the pressure minimum are determined numerically (Fig. 1). Here, the search of the minimum line is made by the use of the first and second derivatives of pressure field so that a numerical data of good quality is desired. Since the low-pressure regions do not always represent swirling motions, a swirl condition is imposed on the lines. The core is defined as such regions in which the pressure is concave in the radial directions, i.e. $\partial^2 p/\partial r^2 > 0$, where r is the local radial

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 3–16, 2001.
© Springer-Verlag Berlin Heidelberg 2001

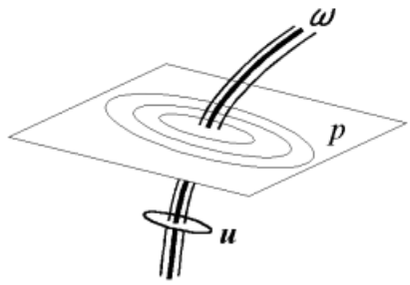

Fig. 1. A low-pressure vortex. The axis is given by a minimum line of pressure field and the core is defined as the region in which the pressure is concave in radial directions $\partial^2 p/\partial r^2 > 0$.

coordinate. The method of construction of the low-pressure vortex is described in [12], [7], [8].

The vortical structures visualized by isosurfaces of vorticity and Laplacian of pressure and by the low-pressure vortex are compared in the next section. The cross-section of a tubular core is analyzed with special attention to the cross-axial component of vorticity in §3. Wrapping of vorticity lines and the resulting spiral structure around a vortex filament are discussed by the use of an asymptotic solution of the Navier-Stokes equation in §4. The role of elementary vortex in energy dissipation is discussed in §5. Finally, section 6 is devoted to the concluding remarks.

2 Low-Pressure Vortex in Turbulence

The method of low-pressure vortex is applied to a forced turbulence which is simulated numerically by solving a set of equations of motion of an incompressible viscous fluid,

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\frac{1}{\rho}\nabla p + \nu\nabla^2\boldsymbol{u}, \tag{1}$$

$$\nabla\cdot\boldsymbol{u} = 0 \tag{2}$$

in a periodic cube with period 2π. Here, $\boldsymbol{u}$ is the velocity, t the time, ρ the (constant $(=1)$) density, p the pressure, and ν the kinematic viscosity. To maintain turbulence activity the amplitude of the Fourier coefficients at small wavenumbers ($|\boldsymbol{k}| \le k_{\mathrm{f}}$) are kept constant in all the time, whereas their phases evolve freely as the equations of motion allow. This is a kind of forcing having no additional control parameters. The spatial derivative is calculated by the Fourier spectral method on $N^3 = 512^3$ grid points. The aliasing interactions are removed by phase shift method and octodecahedron truncation [13]. The maximum wavenumber $k_{\max}$ retained is $\frac{1}{2}N = 256$. The time scale of energy-containing eddies may be estimated as $\mathcal{E}/\epsilon = 4.8$, where $\mathcal{E} = \frac{1}{2}\left\langle|\boldsymbol{u}|^2\right\rangle$ is the kinetic energy per unit mass and $\epsilon = \nu\left\langle|\boldsymbol{\omega}|^2\right\rangle$ is the energy dissipation rate, $\boldsymbol{\omega} = \nabla\times\boldsymbol{u}$ is the vorticity, and

Table 1. Simulation parameters

N	$k_{\rm f}$	Re_λ	$k_{\max}/k_K$	$l_K/\Delta x$	$\lambda/\Delta x$	$\mathcal{E}/\epsilon$	Δt
512	$\sqrt{8}$	119	2.8	0.92	20.6	4.8	0.00125

$\langle\ \ \rangle$ denotes the spatial average. The temporal integration is performed by the Runge-Kutta-Gill scheme with time increment $\Delta t = 0.00125$. The simulation was performed over several eddy-turnover times until the statistical equilibrium state is reached. The numerical parameters and the statistical quantities calculated in the equilibrium state are listed in table 1. Here,

$$Re_\lambda = \frac{u'\lambda}{\nu} \tag{3}$$

is the Reynolds number, $u' = \sqrt{2\mathcal{E}/3}$ the rms of a single component of velocity, $\lambda = \sqrt{5\mathcal{E}/\mathcal{Q}}$ the Taylor length, $\mathcal{Q} = \left\langle |\boldsymbol{\omega}|^2 \right\rangle$ the enstrophy, $l_K = (\nu^3/\epsilon)^{1/4}$ the Kolmogorov length, $k_K = 1/l_K$ the Kolmogorov wavenumber, and $\Delta x = 2\pi/N$ is the grid width.

Note that the maximum wavenumber $k_{\max}$ is sufficiently larger than the Kolmogorov wavenumber k_K so that the smallest scale of motions are resolved well. This is important to detect the vortex axes accurately by the present tracking scheme of the loci of sectionally local minimum lines of pressure in which the second derivatives of the pressure field are employed. In fact, as shown in Fig. 2, the present simulation resolves deep the dissipation range. Here, $\mathcal{Q}(k) = k^2 E(k)$ is the enstrophy spectrum.

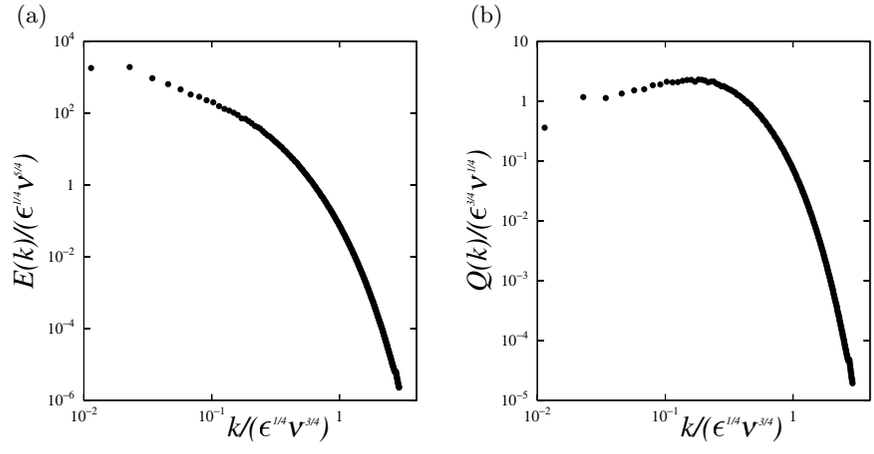

Fig. 2. The normalized energy and enstrophy spectra. The energy dissipation range is well resolved.

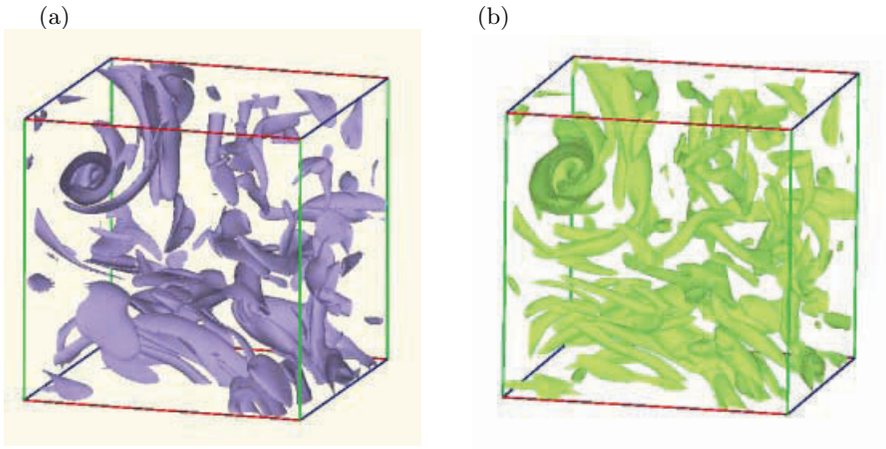

Fig. 3. Isosurface representations of vortical structures. (a) Vorticity magnitude at $|\boldsymbol{\omega}| = 3.7\omega_{\rm rms}$. (b) The Laplacian of pressure at $\nabla^2 p = 3.0\omega_{\rm rms}^2$.

Before describing the pressure method, we present two typical isosurface representations in Fig. 3, where the isosurfaces of magnitude of (a) vorticity and (b) Laplacian of pressure in a 1/64 cube ($0 \le x, y, z \le \frac{1}{2}\pi$) of the periodic box are drawn. The vorticity, by definition, expresses twice of the angular velocity of rotations of fluid elements, whereas the Laplacian of pressure is written as the difference between the symmetric and antisymmetric parts of the velocity gradient tensor,

$$\nabla^2 p = \omega_{ij}\omega_{ij} - s_{ij}s_{ij}, \tag{4}$$

where

$$s_{ij} = \frac{1}{2}\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right), \qquad \omega_{ij} = \frac{1}{2}\left(\frac{\partial u_j}{\partial x_i} - \frac{\partial u_i}{\partial x_j}\right). \tag{5}$$

At the first sight, the shape of isosurfaces is similar between the two (although they change very much depending on the threshold values). The detailed structure, however, is quite different. The vorticity exhibits layer structures, whereas the Laplacian of pressure tubular structures, as has been already noted before (see [14]).

The axes of low-pressure vortices identified by the method of low-pressure vortex are drawn in Fig. 4(a) in the same domain as in Figs. 3. There are 925 vortices identified. It is checked by superposing Fig. 4(a) and Fig. 3(b) that each isosurface of the Laplacian of pressure includes a vortex axis, whereas there are many axes located outside of the isosurfaces. All of these vortices have swirling motions around them (see Fig. 6(e)) [12], but the vorticity on the axis is not necessarily intense. This latter characteristics is a contrast to other vortex tracking schemes of strong vortices (e.g. [3]). One advantage of the present method is a possibility to catch an embryo vortex. In Fig. 4(b) a single axis is picked up with a square panel which is orthogonal to the axis and on which the

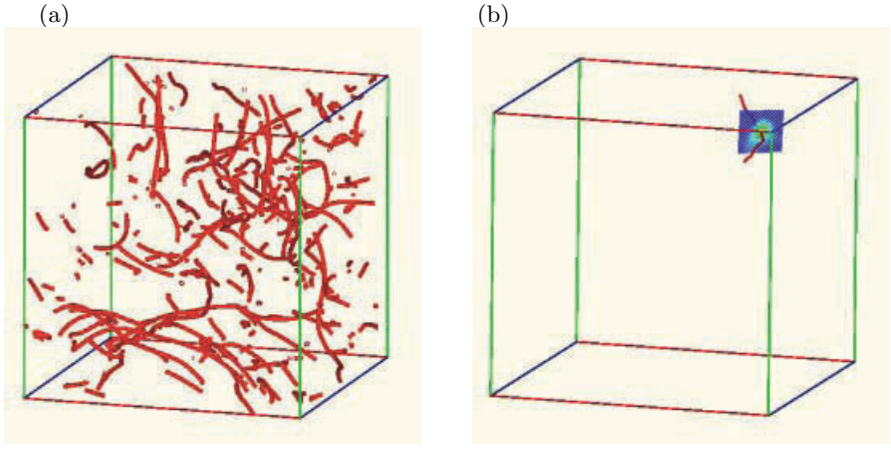

Fig. 4. Axes of low-pressure vortices. (a) 925 vortices are detected in a 1/64 cube (128^3 grid points) of the whole simulation box. The longest vortex is $131\Delta x$ in length. (b) A particular vortex which is analyzed in §3. A square panel with distribution of the Laplacian of pressure is the cross-section on which various fields are drawn in Fig. 6.

distribution of the Laplacian of pressure is shown. This particular vortex line is analyzed in the next section.

The volume occupied by the low-pressure vortex is shown in Fig. 5. Here, the boundaries of all the vortex cores are drawn in (a) and those of 20 longest ones in (b). The shape of the cross-section is mostly round (see Fig. 7(a)). The shape and size of the cores vary along the axis. The mean and standard deviation of the diameter are $12l_K$ and $7l_K$, respectively. The maximum reaches to $71l_K$.

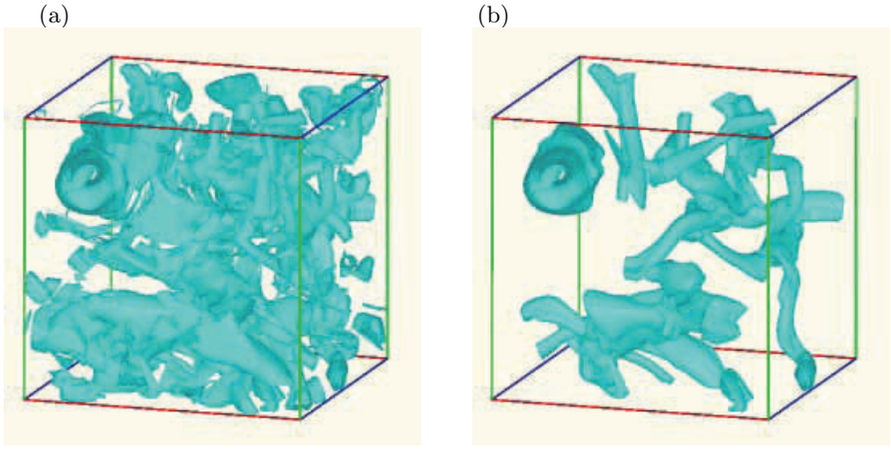

Fig. 5. Cores of low-pressure vortices. (a) The whole vortices occupy 34.9% of the total volume. (b) The cores of the 20 longest vortices. Tubular structures are clearly seen.

3 Structure of Vortex Cross-Section

Tubular concentrated vorticity induces a swirling motion around it. This shear flow stretches vorticity in the direction orthogonal to the vortex tube and generates vortex layers of cross-axial component of vorticity. Such wrapped vortex layers have been observed in flows with fixed directions such as in channel flows (e.g. [1]). But it is difficult to identify them in isotropic turbulence because the decomposition of the axial and cross-axial components cannot be performed unless the direction of a reference vortex is specified which changes in time. It is possible with the present method of low-pressure vortex because the position and direction of each node of all the vortices are known.

In Fig. 6, we plot the cross-sections of various fields on a square plane normal to a vortex axis shown in Fig. 4(b). The side-length of the square is $20\Delta x$ $(= 21.7 l_K)$. The vortex axis is located at the center. Contours of the magnitude, axial and cross-axial components of vorticity are drawn in (a), (b) and (c), respectively. Darker areas indicate larger values. Note that the levels in (a) and (b) are common, but those in (c) are smaller. This means that the intensity of the cross-axial component is quite small compared with the axial component. Nevertheless, a pair of layers of cross-axial component are prominent in (c). This is a variant of the double spirals observed in a decaying turbulence [9].

The periphery of the core boundary, defined by the inflection surface on which the second derivative of pressure with respect to the radial coordinate vanishes for the first time, is plotted in (d). It is seen that the diameter is about 10 times the Kolmogorov length. The spatial distribution of the cross-axial component of velocity relative to that on the axis is depicted in (e). A counter-clockwise swirling motion is evident. The vorticity vector field in the cross-section is shown in (f) The direction of vorticity vectors in the two layers is opposite, which suggests that the origin of these layers may be same, which is in contrast to the case considered in [4] in which they are parallel. Figure (g) confirms that the pressure is indeed lower at the vortex core. The Laplacian of pressure in (h), on the other hand, takes larger values in the core. Actually, as shown in Fig. 3(b) and Fig. 4(b), this vortex is captured as the regions of high-Laplacian of pressure. Finally, the energy dissipation field is shown in (i). As already quoted by many authors (e.g. [10]), most of energy dissipation is taking place at the periphery of vortices and not inside the core (see Fig. 9).

In order to see the vortical structure along a vortex we plot in Fig. 7 the cross-sections of (a) core boundary, (b) the axial and (c) cross-axial components of vorticity, and (d) the dissipation fields at consecutive nodes separated each other by approximately three grid sizes $3\Delta x$. The whole figures cover $12\Delta x$. The cross-section in the third row is the ones shown in Fig. 6. The length scale over which the two-dimensional structure extends is of the order of the vortex diameter. The tubular core surrounded by vortex layers of cross-axial vorticity is a typical vortical structure in turbulence, which we call here the elementary vortex.

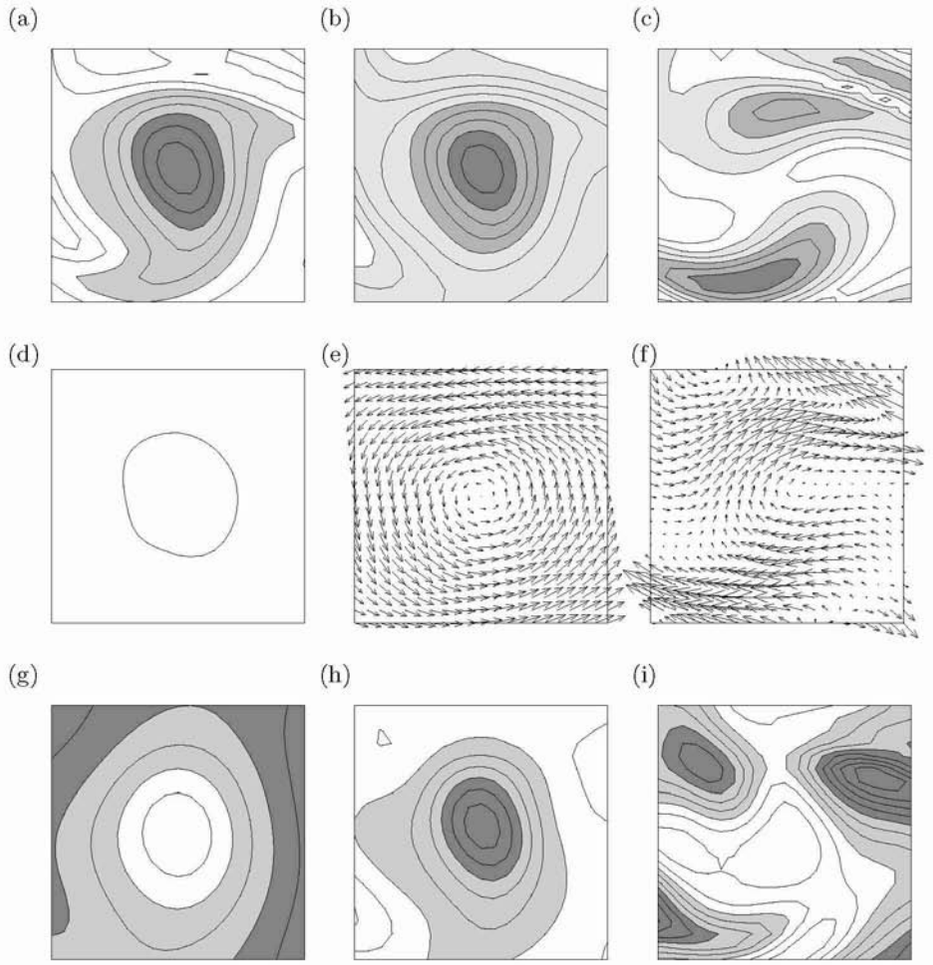

Fig. 6. Various fields across a low-pressure vortex. (a) Vorticity magnitude, (b) axial component of vorticity, (c) cross-axial component of vorticity, (d) periphery of vortex core, (e) cross-axial component of velocity relative to that on the axis, (f) cross-axial component of vorticity, (g) pressure, (h) Laplacian of pressure, and (i) dissipation fields. The contour levels are taken linearly. Darker areas indicate larger values. They are $2, 4, \cdots, 16$ and grayed up at 6 and 12 in (a), $0, 2, \cdots, 16$ and grayed up at 0, 8 and 14 in (b), $1, 2, \cdots, 8$ and grayed up at 2, 4 and 6 in (c). These square boxes are the one shown in Fig. 4(b), the side length of which is $20\Delta x$. The vortex axis is located on the center.

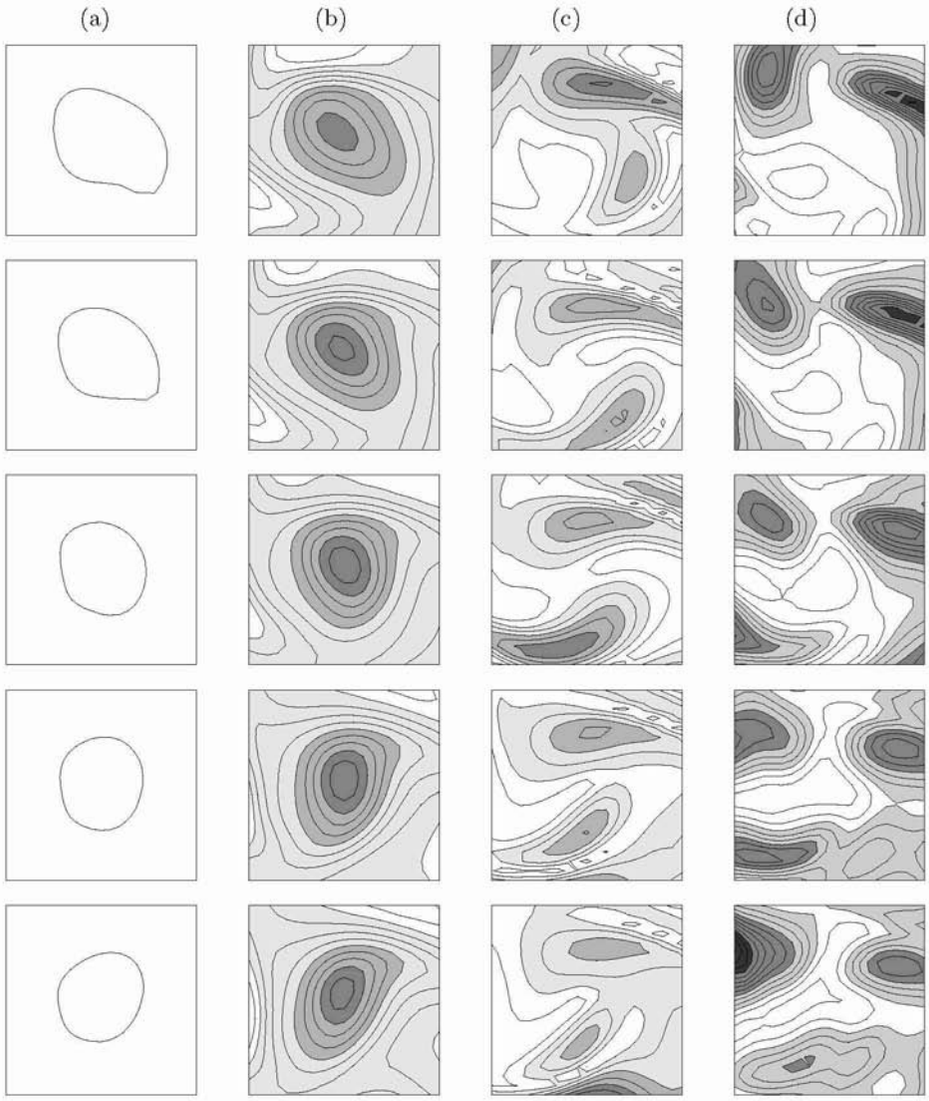

Fig. 7. Variation of cross fields along a vortex. (a) Periphery of vortex core, (b) axial component of vorticity, (c) cross-axial component of vorticity, (d) dissipation field. The four fields in a same row are cross-sections on common planes which are separated each other approximately by $3\Delta x$. The third row corresponds to Fig. 6. The contour levels in (b), (c) and (d) are taken linearly. Darker regions represent larger values.

4 Formation of Vortex Layer

The formation of a vortex layer around a straight vortex filament in a uniform shear flow is described in detail in [4] and [6] in the general case in which the directions of the vortex filament and the shear flow are arbitrary. Here, we recapitulate it in a special case that the vortex and shear vorticity are orthogonal to each other.

Suppose that at the initial instant there are a straight vortex filament on the z axis and uniform vorticity $\boldsymbol{\omega} = S\widehat{\boldsymbol{x}}$ parallel to the x axis (Fig. 8(a)). In the cylindrical polar coordinate system (r, θ, z), there is no radial component of velocity $u_r(r, \theta; t) = 0$, and the velocity field is written as

$$\boldsymbol{u} = -\frac{\Gamma}{2\pi r}\widehat{\boldsymbol{\theta}} + u_z\widehat{\boldsymbol{z}} \tag{6}$$

at all the subsequent times. The flow is uniform in z, i.e. $\partial/\partial z = 0$. Note that we are restricting ourselves in the region away from the vortex filament where the effects of viscous diffusion are neglected. The radial and azimuthal components of vorticity are then expressed as

$$\omega_r = \frac{1}{r}\frac{\partial u_z}{\partial \theta}, \tag{7}$$

$$\omega_\theta = -\frac{\partial u_z}{\partial r}. \tag{8}$$

This implies that the z component of velocity u_z is regarded as the cross-vorticity line, i.e. a line $u_z = \text{const.}$ is parallel to the cross-axial component of vorticity. The z component of vorticity ω_z is negligible.

The z component of the Navier-Stokes equation (1) is

$$\frac{\partial u_z}{\partial t} + \frac{u_\theta}{r}\frac{\partial u_z}{\partial \theta} = \nu\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\right)u_z. \tag{9}$$

The initial uniform vorticity is written as

$$\omega_r(r, \theta; 0) = S\,\cos\theta, \tag{10}$$

$$\omega_\theta(r, \theta, 0) = S\,\sin\theta, \tag{11}$$

and, from (7) and (8), the z component of the initial velocity is

$$u_z(r, \theta; 0) = Sr\,\sin\theta. \tag{12}$$

First, we consider the inviscid case, in which (9) is reduced to

$$\frac{\partial u_z}{\partial t} - \frac{\Gamma}{2\pi r^2}\frac{\partial u_z}{\partial \theta} = 0. \tag{13}$$

The solution of this equation under the initial condition (12) is given by

$$u_z = Sr\,\sin\left(\theta + \frac{\Gamma t}{2\pi r^2}\right). \tag{14}$$

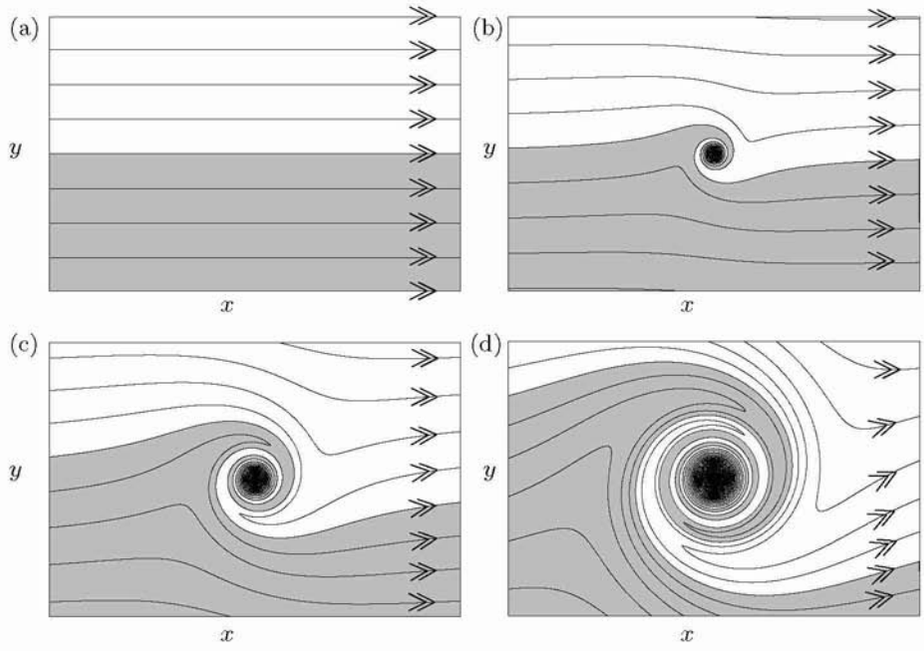

Fig. 8. Wrapping of vorticity lines. The uniform vorticity lines denoted by double arrows are being wrapped by a straight vortex filament which is located at the center and is pointed into the paper. Lines are the contours of $u_z = Sr\,\sin\left(\theta + \Gamma t/2\pi r^2\right)$. It is negative ($u_z < 0$) in gray regions and positive ($u_z > 0$) in white. (a) $\Gamma t = 0$, (b) $\Gamma t = 1/16$, (c) $\Gamma t = 1/4$, (d) $\Gamma t = 1$. $\nu = 0$. (by M. Tanaka)

A cross-vorticity line ($u_z = \text{const.}$) forms a spiral in the region where $\Gamma t/(2\pi r^2) \gg 1$, at which the pitch of the spiral is given by $\Delta r = 2\pi^2 r^3/\Gamma t$.

Next, the effects of small viscosity are taken into account. Guided by the form of (14), we seek the solution in the form,

$$u_z(r,\theta;t) = f(r,\theta;t) Sr\,\sin\left(\theta + \frac{\Gamma t}{2\pi r^2}\right) \tag{15}$$

with the initial condition $f(t,\theta;0) = 1$. Substituting (15) in (9), we obtain

$$\frac{\partial f}{\partial t} - \frac{\Gamma}{2\pi r^2}\frac{\partial f}{\partial \theta} = -\nu\left(\frac{\Gamma t}{\pi r^3}\right)^2 f \tag{16}$$

as long as $\Gamma t/(2\pi r^2) \gg 1$. The solution of this equation is given by

$$f(r,\theta;t) = \exp\left[-\frac{\nu\Gamma^2 t^3}{3\pi^2 r^6}\right] \tag{17}$$

Therefore, the entire solution (15) is written as

$$u_z = Sr\,\exp\left[-\frac{\nu\Gamma^2 t^3}{3\pi^2 r^6}\right]\sin\left(\theta + \frac{\Gamma t}{2\pi r^2}\right). \tag{18}$$

The vorticity is dominated by the θ component,

$$\omega_\theta = S\frac{\Gamma t}{\pi r^2}\left(\frac{\Gamma t}{2\pi r^2}\right)^3 \exp\left[-\frac{\nu\Gamma^2 t^3}{3\pi^2 r^6}\right]\cos\left(\theta + \frac{\Gamma t}{2\pi r^2}\right). \tag{19}$$

This has the same θ dependence as the inviscid solution (14), so is the same spiral structure. The effects of viscosity are reflected in the strong r dependence of the vorticity magnitude. It vanishes both at small and large values of r and has a single maximum

$$\omega_{\theta\max} = 2S\left(\frac{\Gamma}{8\pi\nu \mathrm{e}}\right)^{1/3} \tag{20}$$

at

$$r_{\max} = (\nu t)^{1/2}\left(\frac{\Gamma}{\pi\nu}\right)^{1/3}, \tag{21}$$

at which the pitch of the spiral is estimated as $\Delta r = 2\pi(\nu t)^{1/2}$.

5 Role of Elementary Vortices

The importance in turbulence dynamics of the elementary vortex is controversial. According to Jiménez & Wray [2], the volume fraction of vortex tubes scales as $V_{\mathrm{core}} \propto {Re_\lambda}^{-2}$ and the vorticity magnitude scales as $\omega_{\mathrm{core}} \propto \omega' {Re_\lambda}^{1/2}$, though the definition of vortices is different from ours. The enstrophy of vortex tubes then scales as $\omega_{\mathrm{core}}^2 V_{\mathrm{core}} \propto \omega'^2 {Re_\lambda}^{-1}$, and so does the energy dissipation. This may lead to the conclusion that the dynamical importance of vortex tubes diminishes as the Reynolds number increases. As will be shown below, however, the surrounding vortex layers accompanied with each tubular core can contribute more to the dissipation.

Suppose that the elementary vortex is composed of a tubular core and surrounding vortex layers, and estimate the contribution to the energy dissipation of each part. Assume that the vorticity magnitude, the diameter and the length of tubular cores scale as

$$\omega_{\mathrm{core}} \propto \omega' {Re_\lambda}^{\alpha_{\mathrm{core}}}, \tag{22}$$

$$l_{\mathrm{core}} \propto LRe_\lambda^{\beta_{\mathrm{core}}}, \tag{23}$$

$$L_{\mathrm{core}} \propto LRe_\lambda^{\gamma_{\mathrm{core}}}, \tag{24}$$

respectively, where ω' is the rms of vorticity fluctuation and L the integral length scale. The exponents α_{core}, β_{core} and γ_{core} are constants to be determined theoretically or experimentally. Then the energy dissipation per unit mass due to the tubular core may be estimated as

$$\begin{aligned}\epsilon_{\mathrm{core}} &\approx \nu\omega_{\mathrm{core}}^2 \frac{l_{\mathrm{core}}^2 L_{\mathrm{core}}}{L^3} \\ &\propto \nu\omega'^2 {Re_\lambda}^{2\alpha_{\mathrm{core}}+2\beta_{\mathrm{core}}+\gamma_{\mathrm{core}}}.\end{aligned} \tag{25}$$

Table 2. Scaling exponents of elementary vortex

$\alpha_{\rm core}$	$\beta_{\rm core}$	$\gamma_{\rm core}$	$\alpha_{\rm sur}$	$\beta_{\rm sur}$	$\gamma_{\rm sur}$
$\frac{1}{2}$	$-\frac{3}{2}$	1	$-\frac{1}{6}$	$-\frac{5}{6}$	1

Similarly, under the scaling assumption that

$$\omega_{\rm sur} \propto \omega' {Re_\lambda}^{\alpha_{\rm sur}}, \tag{26}$$

$$l_{\rm sur} \propto LRe_\lambda^{\beta_{\rm sur}}, \tag{27}$$

$$L_{\rm sur} \propto LRe_\lambda^{\gamma_{\rm sur}}, \tag{28}$$

we may estimate the contribution to the energy dissipation of the surrounding vortex layers as

$$\begin{aligned} \epsilon_{\rm sur} &\approx \nu \omega_{\rm sur}^2 \frac{l_{\rm sur}^2 L_{\rm sur}}{L^3} \\ &\propto \nu \omega'^2 {Re_\lambda}^{2\alpha_{\rm sur}+2\beta_{\rm sur}+\gamma_{\rm sur}}. \end{aligned} \tag{29}$$

Since the scaling exponents themselves determine the behavior in the asymptotic limit of large Reynolds numbers, the accurate estimation of them is crucial to discuss the importance of the elementary vortex in this limit. The numerical results by Jiménez et al. [2] [3] for turbulence of Reynolds numbers in the range $37 < Re_\lambda < 142$ suggest the values listed in table 2 for tubular cores as well as the scaling law

$$\frac{\Gamma}{\nu} \propto {Re_\lambda}^{1/2} \tag{30}$$

of the vortex Reynolds number. For the surrounding vortex layers, on the other hand, the exponents may be estimated by taking the rms ω' of vorticity fluctuation for the background strain rate S, and the eddy-turnover time (or life-time of vortex tube) L/u' ($\approx \mathcal{E}/\epsilon$) for the characteristic time t in (20) and (21). The results are listed in table 2. The same scaling as that for the vortex core is employed for the length along the tubular core of the layers.

Substituting these exponents in (25) and (29), we find

$$\epsilon_{\rm core} \propto \nu \omega'^2 {Re_\lambda}^{-1}, \tag{31}$$

$$\epsilon_{\rm sur} \propto \nu \omega'^2 {Re_\lambda}^{-1/3}. \tag{32}$$

for the contribution to energy dissipation of each part of the core and surroundings. According to this estimation both contributions decrease as the Reynolds number increases though the contribution of the surrounding vortex layers is much larger than that of the cores. Since, however, in a view of the numerical as well as theoretical uncertainty of the exponents listed in table 2, it may be premature to draw a conclusion on the limiting behavior.

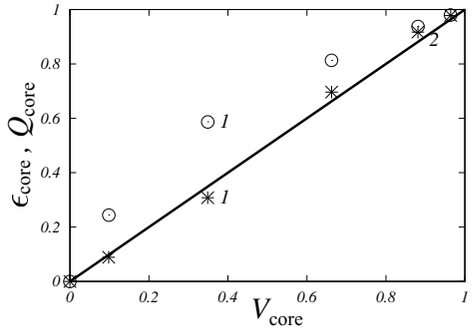

Fig. 9. Contribution of vortex cores to the energy-dissipation and enstrophy. $\epsilon_{\rm core}$ $(*)$, $\mathcal{Q}_{\rm core}$ $(\circ)$, and $V_{\rm core}$ represent the contribution rates from vortex core to the energy dissipation, enstrophy and volume, respectively. Contributions from domains of radii of 0.5, 1, 1.5, 2, and 2.5 multiples of core radius centered vortex axes are shown by symbols. The numbers 1 and 2 attached indicate these multiples.

Although the core regions are defined in the present method, it is not easy to identify the surrounding vortex layers and to calculate the statistics of them. Nevertheless, the data of the coordinates and core radius at each node of all the vortex axes enables us to identify such grid points that are within some distance in the unit of core radius from the axes. The contributions to the energy dissipation ($\epsilon_{\rm core}$, $*$) and enstrophy ($\mathcal{Q}_{\rm core}$, $\circ$) of these inner grid points and the corresponding volume are shown in Fig. 9. Here, the results for the distance of 0.5, 1, 1.5, 2 and 2.5 times the core radius are plotted. The numbers 1 and 2 attached to the symbols denote these multiples. The straight line represents the relation between the volume of the inner regions and their contributions to the energy dissipation and enstrophy if the flow field were absolutely random without any structures. The vortex core occupies 35% of volume, includes 59% of enstrophy and 31% of energy dissipation. Recall that the magnitude of the enstrophy is not taken into account in the process of identification of vortex. Nevertheless, nearly double of the enstrophy is included in the core. On the other hand, the contribution to dissipation is nearly proportional to the volume though less contribution from the inside core is indeed visible. One interesting observation is that the regions within double the core radius from the vortex axes occupy 88% of volume, include 94% of enstrophy and 92% of energy dissipation. This means that the elementary vortex, which is composed of tubular cores and surrounding vortex layers, occupy most of the flow field.

6 Concluding Remarks

The method of low-pressure vortex exhibits new aspects of the vortical structure. By decomposing the axial and cross-axial components of vorticity relative to a vortex axis, we discover the existence of surrounding vortex layers wrapping the tubular core. This layer occupies wider domains and contribute more than the

core part to the energy dissipation. Since the direction of vorticity is orthogonal between the tubular core and surroundings vortex layers, it is expected to play an important role in turbulent mixing and diffusion.

In the present study, only a small portion of a single run is analyzed. In order to get a better statistics it is necessary to gather more data. The Reynolds number dependence of various quantities, such as circulation, size, length, number density of vortices, should be examined to discuss the importance of the elementary vortex in the limit of large Reynolds numbers. We are currently analyzing numerical data of Reynolds numbers 81 ($k_{\max}/k_K = 5.1$), 119 ($k_{\max}/k_K = 2.8$, present study), and 171 ($k_{\max}/k_K = 1.7$) with resolution 512^3. More definite Reynolds number dependence of various statistics is expected to be obtained.

References

1. J.Jiménez, P. Moin: J. Fluid Mech. **225** (1991) 213.
2. J.Jiménez, A.A. Wray: J. Fluid Mech. **373** (1998) 255.
3. J.Jiménez, A.A. Wray, P.G. Saffman, R.S. Rogallo: J. Fluid Mech. **255** (1993) 65.
4. G. Kawahara, S. Kida, M. Tanaka, S. Yanase: J. Fluid Mech. **353** (1997) 115.
5. G. Kawahara, S. Kida, S. Yanase, M. Tanaka: 'Energy dissipation in spiral vortex layers wrapped around a diffusing straight vortex tube'. In: *EUROMECH 8th European Turbulence Conference at Barcelona, Spain, June 27-30, 2000*
6. S. Kida, G. Kawahara, M. Tanaka, S. Yanase: 'Vorticity dynamics around a straight vortex tube in a simple shear flow'. In: *IUTAM Symp. on Dynamics of Slender Vortices at Aachen, Germany, 31 Aug. - 3 Sept. 1997*
7. S. Kida, H. Miura: Eur. J. Mech. B/Fluids **17** (1998) 471.
8. S. Kida, H. Miura: J. Phys. Soc. Jpn. **67** (1998) 2166.
9. S. Kida, H. Miura, T. Adachi: 'Spatiotemporal structure of tubular vortices in turbulence'. In: *IUTAM Symposium on Geometry and Statistics of Turbulence at Kanagawa, Japan, Nov. 1-5, 1999*
10. S. Kida, K. Ohkitani: Phys. Fluids A **4** (1992) 1018.
11. T. S. Lundgren: Phys. Fluids **25** (1982) 2193.
12. H. Miura, S. Kida: J. Phys. Soc. Jpn. **66** (1997) 1331.
13. S.A. Orszag: Stud. Appl. Math. **50** (1971) 293.
14. M. Tanaka, S. Kida: Phys. Fluids A **5** (1993) 2079.

Self-Sustaining Mechanisms of Wall Turbulence

Alexander J. Smits[1] and Carl Delo[2]

[1] Department of Mechanical and Aerospace Engineering, Princeton University, Princeton NJ 08544, USA
[2] NASA Goddard Institute for Space Studies, Columbia University, New York, New York, 10025

Abstract. A general description of the role of organized motions in wall-bounded turbulence is given, with an emphasis on self-sustaining mechanisms. In addition, some proposals are made regarding future directions of research.

1 Introduction

Early turbulence research was based on the statistical interpretation of the velocity field. The experiments and theories were focused on the behavior of the mean flow, the moments of the probability density distributions, and spectra. About 35 years ago, turbulence research developed a new interest in the instantaneous and ensemble-averaged behavior of the turbulent motions themselves. The statistical approach did not lose its vigor, but it was complemented by a growing interest in turbulence structure. This new emphasis was driven, at least in part, by the hope that a closer inspection of the dynamics of turbulent motions would provide the understanding that was needed to develop closure models for the Reynolds-averaged equations. It also generated some exciting observations on the structure of turbulence, especially for wall-bounded flows, that led to the concepts of coherent motions, the burst-sweep cycle, and new scaling laws for the inner and outer turbulence variation with Reynolds number.

The collection of papers recently published in *Self-Sustaining Mechanisms of Wall Turbulence* [1] provides a timely review of the progress in the study of turbulence structure. Many of the principal researchers working in this area contributed a chapter, although the collection is largely focused on the experimental evidence and the effects of compressibility and heat transfer are not considered. It provides, in one volume, most of the current thinking on the structure of wall-bounded turbulence in canonical flows, and it is a must-read for anyone working in turbulence.

Several observations stand out: (1) The contribution of structural models to prediction methods is still in its infancy, although structural thinking has become part of the fabric of turbulence research; (2) Almost all of the work on structure is limited to kinematical rather than dynamical descriptions; (3) A consensus exists among workers on the main structural features of wall-bounded turbulence, despite their statements to the contrary; and (4) There is no clear agreement on the path forward.

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 17–38, 2001.
© Springer-Verlag Berlin Heidelberg 2001

In this paper, we will try to summarize the consensus that emerges from the views expressed in this volume as well as other recent work, and propose some directions for the path forward.

We begin with a short discussion on scaling concepts, and an overview of boundary layer structure. We will also try to point out some of the neglected evidence and newer concepts as we go.

2 Scaling of the Mean Flow

For wall-bounded turbulent shear flows, the velocity field is usually divided into an inner and an outer region. For each region, a length and velocity scale may be defined. The velocity scale in the near-wall region is typically taken to be the friction velocity. The length scale associated with the inner region is then the kinematic viscosity ν divided by the friction velocity, ν/u_τ. For the outer region, the velocity scale is also typically taken to be the friction velocity, and the length scale is taken to be the radius of the pipe R or the boundary layer thickness δ.

The scaling is important, in that it connects laboratory experiments peformed at low Reynolds number to higher Reynolds number typical of most practical applications. It also helps to identify the source of a coherent motion. For example, near-wall motions are expected to scale with inner-layer variables. If we observe that, as the Reynolds number changes, some part of the inner-layer structure depends on outer-layer variables, we would logically conclude that this is evidence for a dynamical connection between the outer and inner layer motions.

Recently, it was suggested by Zagarola & Smits [2], [3] that the outer velocity scale should be independent of the wall stress. By examining the mean velocity profiles of fully-developed pipe flows and turbulent boundary layer flows over a large Reynolds number range, they suggested a "true" outer velocity scale u_0, where $u_0 = U_{CL} - \bar{U}$ for pipe flow and $u_0 = U_\infty \frac{\delta^*}{\delta}$ for boundary layers, where U_{CL} is the centerline velocity for a pipe, U_∞ is the freestream velocity for a boundary layer, $\bar{U}$ is the average velocity, and δ^* is the displacement thickness. The implication may be that some conclusions regarding the interaction between inner and outer layer motions may need to be re-examined in light of these new scaling laws. However, since u_τ has been almost universally used as the outer velocity scale, it will take a great deal of effort to recast the data in terms of u_0. In this paper, we will need to accept previous conclusions regarding the scaling of the outer layer, but it is possible that this assumption may lead to incorrect conclusions regarding the dynamics of the boundary layer.

Although we are mostly concerned with boundary layers, the bulk of the evidence for inner-layer behavior is based on channel flow or pipe flow data. However, comparisons between boundary layers and pipe flows must be made very carefully. Even though a similar scaling may exist, we can not expect the functional forms to be the same since the particular equations of motion and boundary conditions are different. This is true even in the infinite Reynolds number limit.

3 Background on Boundary Layer Structure

Any discussion of the spatial organization of wall-bounded turbulent flows requires a definition of what we mean by "coherent structure." The terms "coherent motion" or "organized motion" are often used as synonyms for coherent structure, although there is a formal difference between a "structure" and a "motion." The most widely-accepted definition is probably that due to Robinson [4]:

> "... a three-dimensional region of the flow over which at least one fundamental flow variable (velocity component, density, temperature, etc.) exhibits significant correlation with itself or with another flow variable over a range of space and/or time that is significantly larger than the smallest local scales of the flow."

As Kline & Robinson [5] point out, the classification of coherent structures in turbulent boundary layers can be arranged in many ways, and they use the term "quasi-coherent" to emphasize that the structures or motions exhibit significant variation from one experiment to another. Eight categories of quasi-coherent structures were listed:

1. Low speed streaks in the near-wall region; $0 \leq y^+ \leq 10$ (see figure 1).
2. Ejections of low-speed fluid outward from the wall.
3. Sweeps of high-speed fluid toward the wall.
4. Near-wall "pockets," observed as areas clear of marked fluid in certain types of visualizations.
5. Strong internal shear layers in the wall zone; $y^+ \leq 80$.
6. Vortical structures of several proposed forms.
7. Large (δ-scale) discontinuities in streamwise velocity.
8. Large-scale motions in the outer layer (including, for boundary layers, bulges, superlayers and deep valleys of free-stream fluid — see figure 1).

Although listed as structures, ejections and sweeps are motions; taken together, they are the main components of the turbulent bursting cycle. All these terms are widely used in [1]. Items 1–5 above are associated with the inner layer. Items 6–8 are considered outer-layer structures, although they may extend across the boundary layer into the near-wall region. In this sense, vortex loop structures (Item 6) and the LSM (Item 8) are the attached eddies hypothesized by Townsend. However, the depth to which such outer-layer structures extend may be a function of Reynolds number.

Cantwell [6] examined the coherent structures in the context of localized regions of vorticity, and suggested instead four main constituents of the organized structure, grouped into inner and outer layer structures as follows:

> **Near-wall structure:** "Nearest the wall is a fluctuating array of streamwise counter-rotating vortices. The vortices densely cover all parts of the smooth wall. Slightly above the streamwise vortices but still quite close to the wall

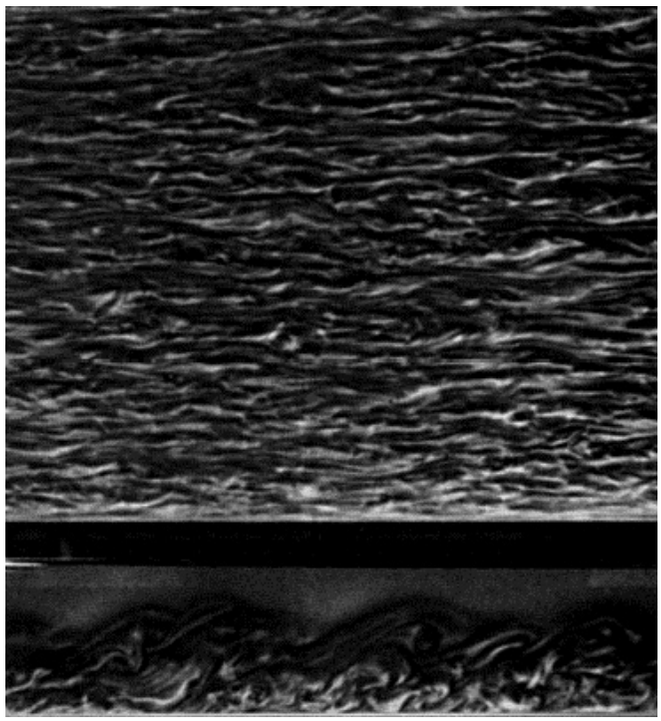

Fig. 1. Turbulent boundary layer on a flat plate. The water was seeded with aluminum particles and viewed simultaneously from above and the side. Note the large population of meandering near-wall streaks in the top view, and the downstream-leaning, large-scale structure in the side view [36].

is a layer that is regularly battered by bursts that involve very intense small-scale motions of energetic fluid. ('energetic near-wall eddies')."

Outer-layer structure: "The outer layer is also occupied by intense small-scale motions ('energetic outer flow eddies'). These are found primarily on the upstream-facing portions of the turbulent/non-turbulent interface; the backs of the bulges in the outer part of the layer. The outer small-scale motions are part of an overall transverse rotation with a scale comparable to the thickness of the layer."

Cantwell states that the low-speed streaks are experimentally observable evidence of pairs of streamwise-oriented, counter-rotating vortices in the near-wall region. Smith *et al.* [7] (also chapter 2 in [1]) present convincing arguments that the low speed streaks are in fact artifacts of the passage of symmetric and asymmetric hairpin vortices. These two views can be reconciled by the assumption that the streamwise vortices in the near-wall region are the trailing "legs" of "hairpin" or "lambda-shaped" or "horse-shoe" vortices that are attached to the wall (see figure 2).

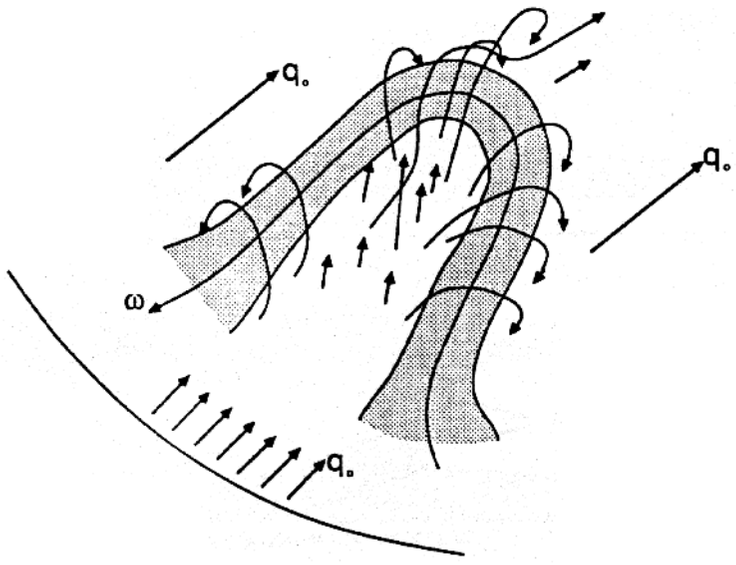

Fig. 2. The inclined horse-shoe vortex first proposed by Theodorson as the basic structure of all turbulent boundary layers.

Ejections, sweeps, pockets and strong internal shear layers (Items 2–5 above) may be interpreted as separate forms of the generic, energetic near-wall eddies proposed by Cantwell, or as forms of the "typical eddies" proposed by Falco (see also Morrison *et al.* [8]). The energetic outer flow eddies are also analogous to Falco's typical eddies, specifically those that reside on the backs of the LSM's in the outer layer (see chapter 7 in [1], and figure 3).

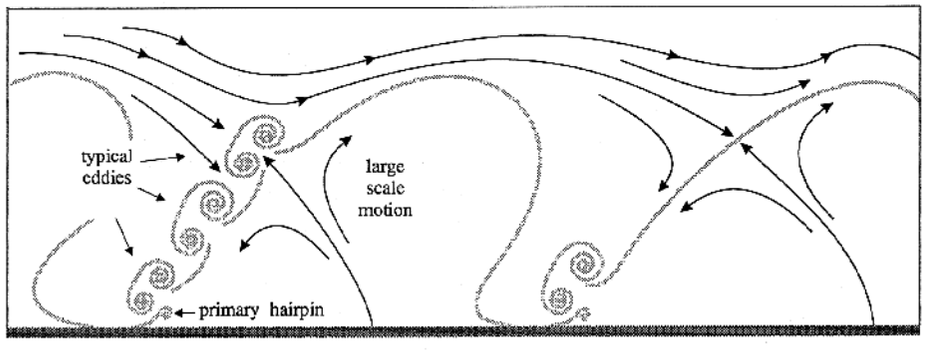

Fig. 3. Schematic representation of typical eddies grouped on the back of a large-scale motion [37].

3.1 Inner Layer Motions

The average spanwise spacing of the streaks ($\lambda^+ = \lambda u_\tau/\nu \approx 100$, where λ is the spanwise wavelength) is one of the few universally accepted aspects of boundary layer structure, although there are large departures from the mean [9]. In contrast, estimates of the streamwise extent of the streaks vary widely. This is most likely due to the different methods used to make the estimates. The low-speed streaks are seen to extend up to $L_x^+ = Ł_x u_\tau/\nu \approx 2000$ in some cases, if the spanwise meandering of the streaks is taken into account (as in visual studies). Velocity correlation methods derived from stationary probes would tend to underestimate the streamwise length of a meandering streamwise vortex, as only limited portions of such a vortex would intersect the probe. In addition, as Cantwell points out, even if the low-speed streaks are generated directly by pairs of counter-rotating vortices, it does not necessarily follow that the two must have the same streamwise dimension (chapter 2 in [1], and [7]). The persistence distance of the near-wall vortices is difficult to evaluate for the same reasons. Probably the most complete description of their formation is given by Schoppa & Hussain (see chapter 16 in [1]).

The spanwise dimension of the near-wall energetic vortices, should be approximately $L_z^+ \approx 50$, based on the observed mean spanwise spacing of the low-speed streaks. The spanwise dimension of the energetic near-wall eddies is not clear, but it is probable that these structures have limited spanwise extent, of the same order as the streamwise dimension, that is, $L_z^+ \approx$ 20-40.

The convection velocities of the various types of near-wall structures is also fairly well established, and the values represent data from a wide range of studies, primarily correlation-based. The average convection velocity of the near-wall streamwise vortices is probably close to the local mean velocity, as Cantwell reports that the low and high speed streaks move at about 1/2 and 3/2 of the local mean velocity respectively.

3.2 Outer Layer Motions

The main emphasis in [1] is on the near-wall motions. This is a natural bias since they are associated with high levels of turbulence production and dissipation, and therefore must feature prominently in any discussion of self-sustaining mechanisms. It is also important to consider the outer layer motions. The LSM's are still an active feature in that up to 40% of the outer-layer Reynolds shear stress can be found in the neighborhood of their sloping interfaces [10]. They are responsible for the large-scale transport of turbulence, the growth of the layer by entrainment of initially irrotational flow, and appear to play a crucial role in triggering instabilities in the near-wall region.

Although the outer-layer motions have not been studied as extensively as the inner layer motions, the general characteristics of the LSM's are well known. The large-scale motions (Townsend's "main eddies") evolve and decay slowly as they convect downstream, and, on average, they are inclined to the wall at an acute angle, leaning in the downstream direction [11]. The most identifiable

feature of these motions is a downstream-sloping shear-layer interface between upstream high-speed fluid and downstream low-speed fluid. These interfaces are three-dimensional shear layers which form the upstream side of the largest of the boundary-layer eddies, and remain coherent long enough to convect several boundary-layer thicknesses downstream.

Between neighboring bulges, the flow is irrotational, resulting in the intermittent character of the outer layer. Figure 4 shows the presence of several large-scale motions or bulges in a boundary layer. The structures are seen to vary greatly in size and inclination angle. The properties of the large-scale motions, such as length scales, time scales, convection velocity, and structure angle, as well their internal structure, such as velocity, vorticity, and pressure fields, remain the subject of controversy and active research. Furthermore, the Reynolds-number-dependence, if any, of the large-scale motions is not very well understood, although recent measurements in the Superpipe experiment at Princeton indicates that no clear high-Reynolds-number limit is evident [12].

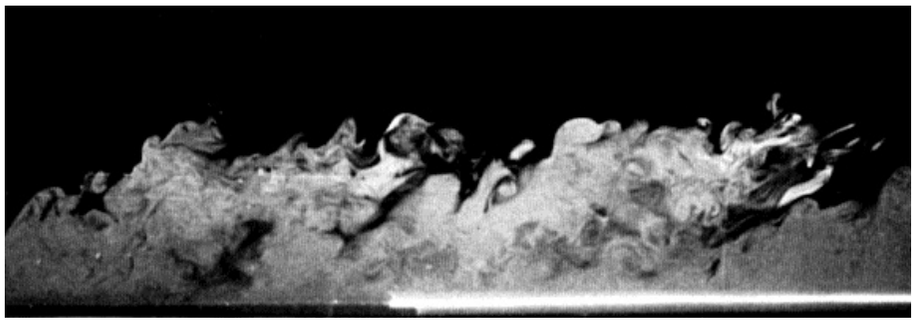

Fig. 4. Flow visualization of a boundary layer in subsonic flow at $R_\theta \approx 4,000$, obtained by seeding the flow with a fog of oil droplets, and illuminating the flow with a planar laser sheet. Flow is from left to right [17].

The flow visualizations by Head & Bandyopadhyay [13] provide strong support for the concept that at least some of the LSM's are loop-shaped vortical structures, and they showed that the aspect ratio of these vortical structures is Reynolds number dependent, in contrast to Townsend's Reynolds number similarity hypothesis. Figure 5 shows that at $R_\theta = 600$ the structures have the proposed horse-shoe shape, with a low aspect ratio. As R_θ increases, the structures become elongated, and at $R_\theta = 9,400$ they appear more as hairpins, with a large aspect ratio. At all Reynolds numbers, the spacing between the legs of the structures is similar to the spacing between the near-wall streaks ($\approx 100\nu/u_\tau$). The structures appear to extend to the wall, in support of the attached eddy hypothesis, although it is difficult to believe that the legs could remain coherent and connected to the heads at such large distances from the wall (in terms of y^+). Head & Bandyopadhyay also suggested that at low Reynolds numbers the LSM's were merely single horse-shoe elements, but at higher Reynolds numbers

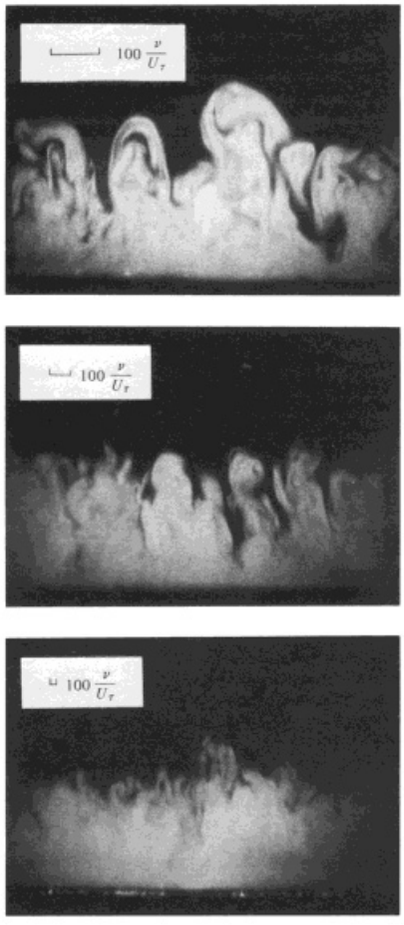

Fig. 5. Flow visualizations by [13], showing the Reynolds-number-dependence of vortex-loop structures in a turbulent boundary layer. The visualizations were obtained by filling the boundary layer with smoke, and illuminating the flow with a laser sheet inclined 45° downstream. a) $R_\theta = 600$; b) $R_\theta = 1,700$; c) $R_\theta = 9,400$ [13].

the LSM's were actually agglomerations of many elongated hairpin vortices. At low Reynolds numbers, the LSM's exhibited a "brisk" overturning motion, while at higher Reynolds numbers they overturned slowly. This suggests that entrainment decreases with increasing Reynolds number, which is in agreement with the observation that the boundary layer grows more slowly with increasing Reynolds number.

Head & Bandyopadhyay also observed the generation of multiple hairpin loops in "packets," where the heads of the loops lie along a line that is inclined to the wall at an angle of about 20°, as shown in figure 6 (see also section 4).

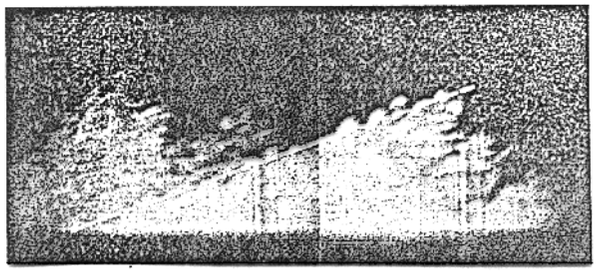

Fig. 6. Example of 20° interface at $R_\theta = 17,500$. Flow is from left to right [13].

3.3 Interactions

Sreenivasan & Sahay (chapter 11 in [1]) describe the Reynolds number dependence of the peak of the Reynolds shear stress in its location (y_p^+) and maximum value. Combined with the Reynolds number dependence of the peak in the streamwise turbulence intensity [14], these observations indicate that the near-wall region is not entirely independent of the outer layer dynamics. It has long been suspected that the inner and outer layer motions interact in a complex way, and the rather inconclusive discussion that has taken place regarding the scaling of the bursting frequency illustrates the problem [6,15].

Nevertheless, some notable observations stand out. For example, Brown & Thomas [16] used space-time correlations and conditional sampling techniques to investigate the relationship between the large-scale motion and the wall shear-stress in a boundary layer at $R_\theta = 4,940$ and $10,160$. They found that large-scale motions were inclined at an angle of 18° to the wall and extended $\approx 2\delta$ in the streamwise direction (recall Head & Banyopadhyay's observations on hairpin packets, figure 6). As the structures passed over the wall, they created a characteristic wall-shear-stress signature. Brown & Thomas suggested that this wall-shear-stress pattern was related to the bursting process, and concluded that the large-scale, outer-layer structure influenced the near-wall structure and dynamics. They found that the results were independent of Reynolds numbers when scaled on outer variables.

In addition, the flow visualizations by Falco [17] showed that there may be at least two types of organized motions in the outer layer: LSM's and "typical eddies" (see also chapter 7 in [1], and the classification by Cantwell). He found that the typical eddies are small-scale motions, which scale on wall variables and are responsible for a significant fraction of the total Reynolds shear stress in the outer layer. The average streamwise extent of the LSM's was about 1.6δ at $R_\theta \approx 1,000$. The streamwise length scale of the typical eddies had a constant value of about $200\nu/u_\tau$ for $1,000 < R_\theta < 10,000$. The vertical length scale varied nearly linearly from $100\nu/u_\tau$ to $150\nu/u_\tau$ over the same range of Reynolds numbers. Falco found that the typical eddies generally appear on the backs of the LSM's, and propagate toward the wall, thus acting as sweeps very near the wall.

Falco claimed that the typical eddies may be an intermediate link between the inner and outer layers. As discussed by [18], however, if the typical eddies scale on wall variables then at very high Reynolds number, where δ^+ is very large, the typical eddies will become vanishingly small compared to the boundary layer thickness, and they are unlikely to be dynamically significant (that is, they will not carry significant levels of shear stress). Furthermore, flow visualizations at very high Reynolds numbers show features which appear to be very similar to the typical eddies observed by Falco at lower Reynolds numbers, but which are at least an order-of-magnitude larger (in terms of inner variables), even when the variations in fluid properties, as expressed by the difference between R_θ and R_{δ_2}, are taken into account [11].[1]

Murlis & Bradshaw [19] concluded that there is a large variation of eddy structure with Reynolds number, but that the basic transport mechanisms do not vary appreciably. They also suggested that the large-scale motions carry more shear stress at higher Reynolds numbers, owing to the diminishing importance of the typical eddies.

Smits *et al.* [20] proposed that as the Reynolds number increases the number of discrete scales of motion also increase. At the lowest Reynolds numbers ($1,000 < R_\theta$ and $\delta^+ < 500$), a single scale dominates. This scale must be the same as that of the typical eddy, since they are of the same order as the boundary layer thickness. At higher Reynolds numbers ($1,000 < R_\theta < 5,000$ and $500 < \delta^+ < 2,500$), two scales appear: the large-scale outer-layer bulges and Falco's typical eddies. The typical eddies still make significant contributions to the Reynolds stresses. With further increase in Reynolds number, a third scale appears, the smaller bulges seen at the edges of the large-scale bulges. The typical eddies are now much smaller than the energy-containing range (which scales with δ), so that they are no longer dynamically important (with $Re_\theta = 35,000$, $200\nu_w/u_\tau$ corresponds to a frequency of about $100U_e/\delta$). It may be expected that the third scale would scale on outer-layer variables, although some form of mixed scaling may also be appropriate. The appearance of new scales with increasing Reynolds numbers at approximately equal intervals in $\log \delta^+$ should be reflected in the behavior of the turbulence, especially the higher-order moments.

3.4 Recent Work on Interactions

The analysis of ensemble-averaged data necessarily disregards the instantaneous global organization of the flow field. As a given portion of a turbulent flow is composed of structures of varying age and level of activity, ensemble-averaged approaches may present a more complicated or confused picture of the flow-field structure than actually exists. The highly complex nature of the three-dimensional, time-evolving velocity and vorticity fields suggests the use of volumetric imaging of a passive scalar to reveal the spatial and temporal characteristics of the large-scale structure. As noted by Brown & Thomas [16] (among others), the essentially three-dimensional character of boundary layer motions

[1] $R_\theta = \rho_e U_e \theta/\nu_e$, and $R_{\delta_2} = \rho_e U_e \theta/\nu_w$

makes unambiguous detection of a complex process such as the turbulence generation cycle from a fixed measurement station difficult. Part of the problem is no doubt due to the fact that more than one fluid packet may be ejected from a given low speed streak, so the identification of a single bursting event becomes ambiguous [21]. Additionally, the different detection methods employed add to the confusion (for example, visual identification, quadrant analysis, conditional sampling, etc.) [22]. In this respect, a complete three-dimensional visualization of the flowfield is more appropriate for revealing the complex spatial interrelationships in the boundary layer.

To this end, Delo [23] and Delo & Smits [24] obtained time resolved three-dimensional visualizations of the scalar field in a boundary layer at $Re_\theta = 701$. The volumes were collected in the form of image stacks; each volume was composed of 20 plan-view images of scalar intensity, uniformly separated in the wall-normal direction.

The experiments were conducted in a free-surface water channel with a full-width flat plate positioned in the test section (figure 7). Disodium fluorescein dye was introduced from two spanwise dye slots, illuminated with an argon ion laser via a laser sheet scanning apparatus. The laser sheets were imaged with a high-speed analog video system, then digitized and reassembled into volumes for subsequent visualization and quantitative analysis.

The two-point flow-tagging scheme developed by Goldstein & Smits [25] was used in order to distinguish between fluid in the inner and outer portions of the boundary layer. The upstream dye slot, situated just behind the trip wire, was used to give a discernible background dye level to the turbulent fluid throughout the boundary layer. The dye was introduced in a heavily concentrated form to offset the effects of turbulent diffusion and the entrainment of un-dyed freestream fluid. The downstream dye slot was placed just upstream of the interrogation volume in order to brightly mark near-wall fluid. Because of the close proximity of the second slot, the dye released was of a lower concentration, while still being distinguishably brighter than the dye marking the outer layer fluid. The scalar released from the first dye slot was assumed to give a good indication of the large-scale structure in the boundary layer, particularly the turbulent-non-turbulent interface, and the scalar released from the second dye slot was tracked in order to investigate near-wall motions (see discussion in [25] and [24] for further details).

The scalar structures observed in the primary experimental volumetric scalar data set (50 volumes) were similar to those seen in numerous experimental investigations: two-dimensional cross sections of the observed three-dimensional scalar structures, when compared to previous two-dimensional flow visualization studies, demonstrated this conclusively. In spite of the relatively small number of scalar structures observed, and the high degree of variability in their size and shape, certain aspects of the structure and dynamics of the scalar field occurred with regularity, and were felt to be representative of turbulent boundary layers in general.

One new aspect of this study was the large-scale spanwise organization of individual scalar structures (streamwise and spanwise scale $\approx \delta$) along diagonal

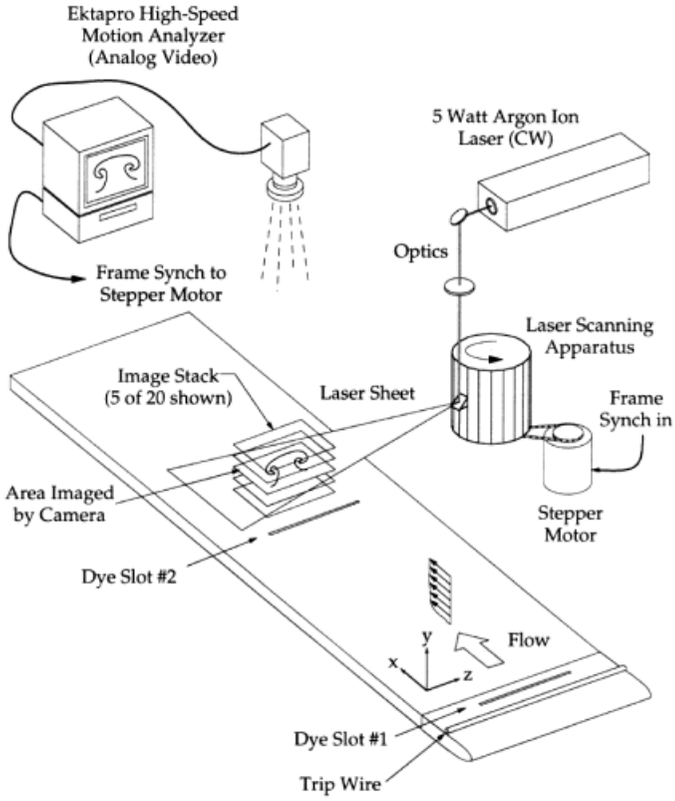

Fig. 7. Experimental configuration for volumetric experiments [24].

lines in the x-z plane. The lines were inclined to the freestream within the range $\pm 50°$. The spacing of the structures along the lines showed significant variability, but it was approximately 1δ.

Especially interesting was the observation that groups of these large-scale scalar structures were frequently assembled into agglomerations measuring up 5δ in length. These very-large-scale, streamwise and spanwise extensive agglomerations were oriented along diagonal lines within the same range noted above ($\pm 50°$). The agglomerations appeared to be spaced far apart in the streamwise direction ($\approx$ 8–10δ). Although relatively few of them were observed because of the large streamwise separation, this estimate was derived from long-time composites of an extended 151 volume data set. The extent of the data volume was not large enough to estimate their distribution in the spanwise direction.

The similar orientation of these very-large-scale agglomerations and the smaller scale, individual structures suggests that the agglomerations are formed through the merging of adjacent, previously existing individual vortical structures as they convect downstream. However, fluctuating streamwise and spanwise

velocities in the mid to outer levels of the boundary layer were found to be too small to bring structures that are separated by $O(\delta)$ together within their lifetime, $O(6\text{–}10\,\delta/U_e)$. Also, vortex merging does not in itself explain the large-scale diagonal organization of the flow in the spanwise plane.

Some evidence was found that the large agglomerations actively "build" themselves as they convect downstream, through the addition of near-wall fluid. Most of the ejection events identified in a spatial distribution occurred simultaneously, soon after the passage of the extended, large-scale structures. The fluid ejected by these events generally penetrated through the trailing edges of the structures, effectively increasing their streamwise extent. It was tentatively proposed that the addition of energetic near-wall fluid in this manner serves to perpetuate the active nature of the large-scale structures. The key seems to be the extended streamwise dimension of the large-scale agglomerations; these structures appear to be "long" enough so that after perturbing the inner layer low-speed streaks and triggering an ejection, the trailing edges of the structures are still in the neighborhood to profit from the addition of the ejected near-wall fluid. Typically, after a streak lift-up and ejection, there is a quiescent period in which no ejections occur, so no growth of scalar structures would occur after the passage of such a large-scale structure. In this way, it may be that the large-scale agglomerations grow preferentially, at the expense of smaller collections of structures.

The general diagonal orientation of the very-large-scale agglomerations and the individual large-scale scalar structures in the outer portion of the boundary layer is similar in some respects to the underlying vortex field of a turbulent spot as proposed by Perry *et al.* [26]. In their model, an initial disturbance to a laminar boundary layer causes three-dimensional undulations in initially straight vortex bundles aligned in the spanwise direction. Fluctuating velocities induced by the wavy lines of vorticity cause the rapid growth of the turbulent spot in both the spanwise and streamwise directions. It is likely that there is a direct connection between the observed growth pattern of turbulent spots and the orientation observed in the turbulent boundary layer.

Another aspect of the large-scale global organization of the flow was the observation that ejections of near-wall fluid appeared to be spatially organized, and related to the passage of the large-scale agglomerations. The motion of the near-wall streaks as they underwent oscillation, lift-up and ejection was reasonably consistent throughout the data set, and in line with previous models. During the ejection process, it was seen that a single streak can be responsible for the ejection of more than one distinct fluid packet. Whether the packets should be considered separate ejections, or components of a single, spatially distributed ejection is unclear, but it seems likely that a connection exists between these multiple ejections and the formation of Head & Bandyopadhyay's hairpin packets.

A further interesting observation was the sinuous spanwise deformation of the streaks during oscillation and lift-up, a clear indication of the three-dimensionality of the process. The highly contorted nature of the streaks, and the wide spatial

distribution of the portions of the streak that were undergoing motion, serve to explain some of the disparate results obtained by researchers concerning the bursting process, particularly the bursting frequency [22], [15].

It was also found that near-wall fluid was sometimes ejected from the near-wall region well into the outer portion of the boundary layer ($y/\delta > 0.7$) while traversing a relatively short streamwise distance. The arrival of fluid with high momentum deficit this far into the outer portion of the boundary layer represents significant momentum exchange. The fact that the near-wall fluid can penetrate this far over a relatively short time span indicates that the outer layer structures are not simply the passive artifacts of events that happened far upstream, at least at the low Reynolds number regime investigated here. At higher Reynolds numbers, the greater separation in scales may serve to shelter the outer layer structure from the influence of the near-wall motions. However, the large-scale spatial organization and near-wall motions which were observed in this low Reynolds number investigation may still be present, although their influence may be confined to the logarithmic region in the boundary layer.

4 Summary View of Boundary Layer Structure

Here, we attempt to provide a unified picture of boundary layer structure. Our aim is to minimize the number of different elements and mechanisms comprising the total. This is in contrast to the eight categories listed by Kline & Robinson [5], with several categories encompassing a number of sub-items. It is proposed here that a conceptual model comprising only a smallm number of elements can provide a sufficient physical basis for a complete description. These elements have counterparts in those identified by Cantwell [6] but we will try to be more specific. This model does not describe all possible observations, but instead tries to represent the important kinematic and dynamic behavior in wall-bounded turbulent flows.

First, we address the question of symmetric and asymmetric vortex loops. The horse-shoe vortex of Theodorson, the hairpin loops of Head & Bandyopadhyay, and the lambda-shaped vortices of Perry & Chong, are symmetric structures. It has been widely reported in experimental studies and Direct Numerical Simulations (DNS) that the vortex structures observed in boundary layers are only rarely symmetric and asymmetric structures dominate (see, for example, [5] and chapter 8 in [1]). Robinson [4] lists quasi-streamwise vortices in the near-wall region, "legs" and "arches" in the overlap region, and arches in the wake region (see figure 8), but even the arches were typically asymmetric. In the three-dimensional visualizations by Delo [23,24], symmetrical structures are rarely seen, although it is readily apparent that legs and arches are always part of the same vortex-loop-like structure. These vortex loops may be twisted and contorted, and certainly show wide variations among themselves, but they can be identified quite clearly as single structures. However, they can only be identified as such by the use of time-resolved three-dimensional visualizations and detailed inspection of each likely motion. Robinson's use of an identification threshold

(based on pressure level in his case) may well have given the impression of a more complex assortment of motions.

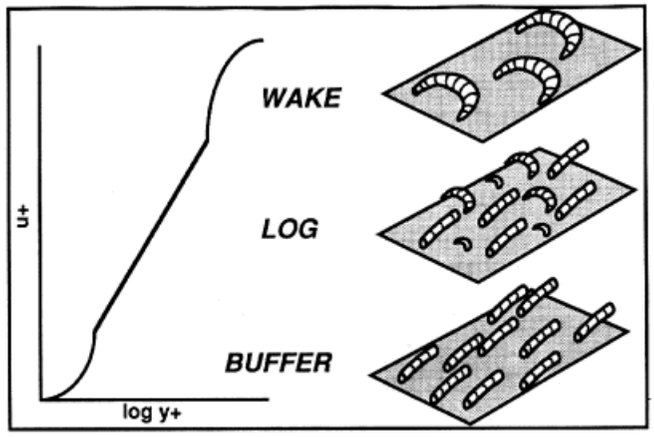

Fig. 8. Idealized schematic of vortical structure populations in the different regions of the canonical boundary layer, according to [4].

Also, as discussed persuasively by Smith & Walker in chapter 2 of [1], any kinematic description of the near-wall cycle in terms of symmetric structures can equally be cast in terms of asymmetric events, even though Klewicki (chapter 7) argues otherwise. Symmetric and asymmetric hairpin vortices are illustrated in figure 9. The important point is the sequence of events and their underlying mechanisms, and since the presence of asymmetry does not seem to affect the underlying physics, symmetrical models can be used for simplicity. In any case, it should be remembered that all descriptions of boundary layer structure represent an idealized model of the behavior, which at best can only capture the average and not the particular.

In that spirit, it seems that the most lucid and complete model of the near-wall kinematics is given by Smith & Walker (chapter 2 in [1]), also [7], [27], [28]), and it provides the basis for the following description, which also incorporates additional observations (see figure 10).

The cycle begins with the growth of a low-speed streak, which continues until the passage of a disturbance of sufficient size and strength impresses a local adverse pressure gradient on the streak causing a local deceleration. The influence of the outer-layer structures is as described by Brown & Thomas [16], Delo [24], and Antonia & Djenidi (chapter 9 of [1]). This deceleration creates a three-dimensional inflectional profile at the interface between the streak and the higher speed flow in the wall region (see Blackwelder in chapter 3 of [1]). The region above the streak is unstable to small local disturbances, and oscillations begin on the top of the streak. The three-dimensional sheet of vorticity above the streak will develop waves due to the oscillations, roll up, and concentrate the

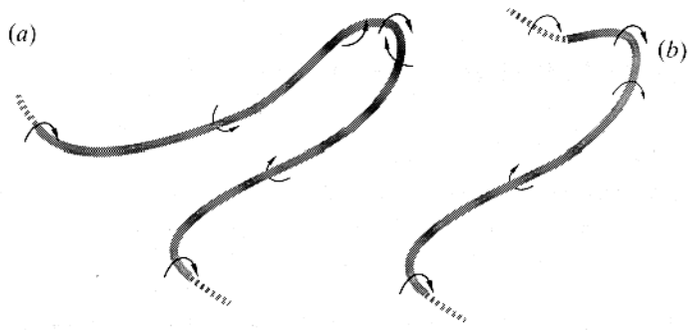

Fig. 9. Schematic diagram of typical hairpin vortex configurations. Sense of vorticity indicated by arrows [7].

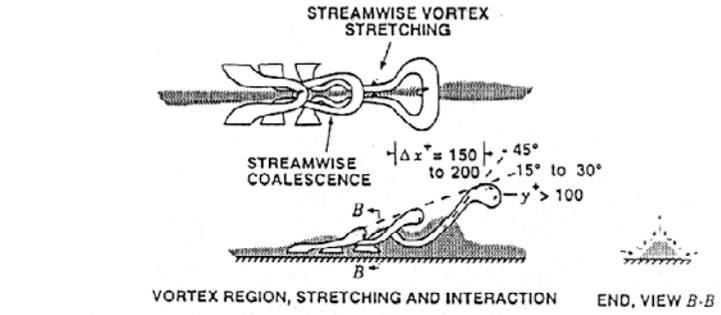
STREAMWISE VORTEX STRETCHING
STREAMWISE COALESCENCE
$\Delta x^+ \equiv 150$ to 200
45°
15° to 30°
$y^+ > 100$
B
B
VORTEX REGION, STRETCHING AND INTERACTION
END, VIEW B-B
A. ILLUSTRATION OF THE BREAKDOWN AND FORMATION OF HAIRPIN VORTICES DURING A STREAK-BURSTING PROCESS. LOW-SPEED STREAK REGIONS INDICATED BY SHADING (FROM SMITH [46]).
OSCILLATION VORTEX ROLL-UP HAIRPIN VORTICES
SYNTHETIC LOW-SPEED STREAK GENERATED DUE TO CONTINUOUS STREAMWISE INJECTION
INRUSH OF FLOW INTENSIFIES WEAK VORTICITY CONCENTRATION REGION
SECONDARY STREAMWISE VORTICES
B. SCHEMATIC OF BREAKUP OF A SYNTHETIC LOW-SPEED STREAK GENERATING HAIRPIN VORTICES. SECONDARY STREAMWISE VORTICAL STRUCTURES ARE GENERATED OWING TO INRUSH OF FLUID (FROM ACARLAR AND SMITH [47]).

Fig. 10. Model for the breakdown and formation of hairpin loops in the near-wall region [27].

vorticity in locations on the sheet. This process takes place at a y^+ of between 15 and 20, and there can be between two and five concentrations per streak.

As the rollup continues, the vortices begin to look like horse-shoe vortices. Biot-Savart interactions between various portions of the vortex amplify the distortion. Self-induced movement away from the wall occurs and the vortex is stretched by the steep mean velocity gradient. Legs of the horse-shoe, which are a pair of counter-rotating vortices oriented in the streamwise direction, develop as the loop is stretched. As the vortex becomes elongated, it begins to look more like a hairpin vortex and the vorticity is increased. The legs tend to pump low momentum fluid away from the wall and thereby perpetuate the low-speed streak. The head or arch creates a streamwise pressure gradient which will cause the liftup and ejection of low-momentum fluid from the low-speed streak as it moves downstream. As the process proceeds, the hairpin continues to be stretched and the head moves farther away from the wall due to self-induction.

As the loops are dragged over the viscous layer near the wall, they help to regenerate new vortices through an interaction with the viscous wall layer (figure 11. The generation of new vortices in this way has been computed using DNS (see Hanratty & Papavassiliou in chapter 5 of [1], and Zhou *et al.* in chapter 6), as well as observed experimentally. The passage of successive loops can tend to focus the streak so that it can continue to grow, or buffet it to weaken it. In either case, the impingement of subsequent vortices create the appearance of waviness and swaying, a commonly observed behavior.

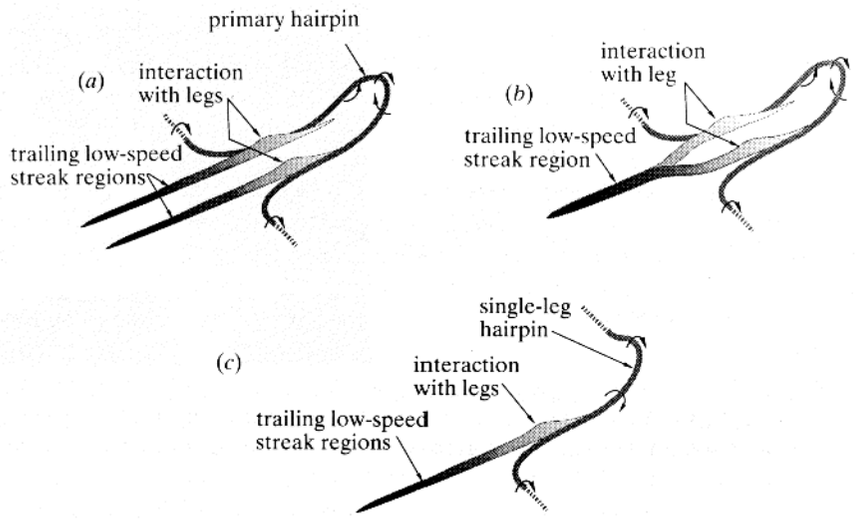

Fig. 11. Schematic illustrating how advecting hairpin-like vortices generate low-speed streaks [7].

The timing on the formation of the loops out of the sublayer material appears to be such that as they rise they form the backs of the LSM's, creating a very active shear layer out of the conjunction of the heads of the vortex loops (see also Bernard & Wallace in chapter 4 of [1], and MacAuley & Gartshore [29]). It was observed by Head & Bandyopadhyay that the hairpins would often align themselves so that their heads lay along a line forming a 15° to 30° angle with the wall, whereas the individual hairpins make an angle of about 45°. A similar picture of hairpin packets emerges from the spectacular PIV measurements by Adrian *et al.* [30].

The stretching of the vortex loops by the surrounding flow (viewed either as a background mean velocity gradient, or the summation of the velocity induced by all surrounding vortex loops) requires work to be done. Also, the narrowing of the vortex tube making up the loop will increase the velocity gradients inside the tube, increasing its dissipation rate as the square of the gradients. This dissipation will be greatest near the wall where the velocity gradients are strongest, and diminish in the outer flow where the gradients are relatively weak. According to Smith & Walker (chapter 2 in [1]):

> ⋯ the process of growth to larger scales in a fully-turbulent boundary layer can be explained by the proximity of multiple hairpin-like vortices in different phases of development, which creates a condition conducive to coalescence, or three-dimensional vortex amalgamation. This amalgamation process ⋯ suggest(s) that local collections of hairpin vortices can intertwine and interact to yield essentially a hairpin-like structure of somewhat larger scale. ⋯ This clearly suggests that the outer region of a turbulent boundary layer can evolve from hairpin vortex structures and that the large, arch-type vortices (Robinson [31]) and rollers (Falco [17]) observed in the outer-region structures directly interface with the freestream flow. The overturning of the structures plays an instrumental part in inducing the flow of higher-speed fluid toward the wall (i.e., intermittent engulfment), which eventually finds its way to the wall along a complex gauntlet of vortex-induced motions, terminating in a "sweep" motion at the wall ⋯ However despite the overall size and strength of the outer structures, the stretching mechanism for energy transfer to these larger-scale vortical structures is strongly diminished because of the weak mean velocity gradient in the outer region. Following an initially strong energy input to the initial vortex scales near the wall, these outer structures will basically "evolve" to larger and larger scales, but with no significant additional energy input, eventually succumbing to slow, viscous dissipation. The outer part of the boundary layer of the boundary layer may thus be regarded as a "graveyard" for vorticity, where the cumulative remnants of deformed wall-region vortices pass through a complicated process of dissipation, diffusion and mutual cancellation, similar to the model of Perry and Chong [32] (figure 12).

The only modifications that should be made to this view are as follows. The heads of the hairpin loops, being a localized concentration of shear stress,

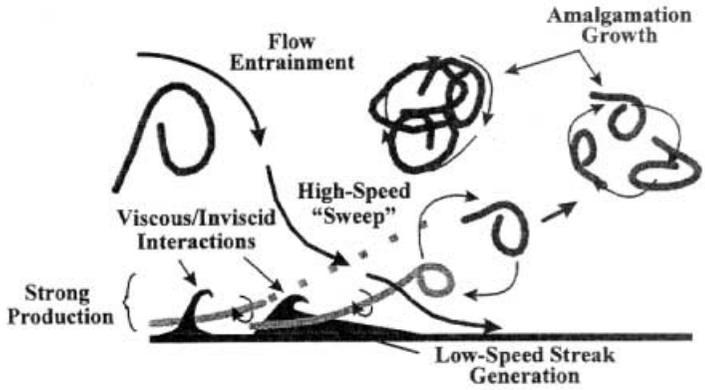

Fig. 12. Conceptual model of turbulence regeneration, amalgamation, and evolution of outer-region structures [7].

continue to be an active feature of the LSM's. In addition, the viscous diffusion of the legs may free the heads to form vortex loops, and at least some of them may have an induced velocity that directs them to the wall, as typical eddies (this view is also advanced by Klewicki in chapter 7 of [1]). As the Reynolds number increases, there needs to be other levels of organization where Very Large Scale Motions emerge since the active motions seen on the backs of the LSM's have a typical dimension of 200 wall units, and these size motions cannot continue to be important in the shear stress budget since the stress in the outer layer scales with outer layer variables.

5 The Path Forward

We have attempted to present a comprehensive view of the kinematics of wall-bounded turbulence. Almost all of the work on turbulence structure is limited to kinematical rather than dynamical descriptions, with two notable exceptions: the work presented by Aubry in chapter 12 in [1], starting from the exceptional contribution by Aubry *et al.* [33], and the newer work of Waleffe & Kim (chapter 13 in [1]). In charting the path forward, it would seem that much more effort needs to be expended on dynamical interpretations. The search for low-order dynamical models may provide the same kind of breakthrough that the horseshoe vortex has provided for the kinematic understanding. Certainly for active and passive control applications such models are essential.

Despite the oft-stated aims of research on structure, the contribution of structural models to prediction methods is still in its infancy. Apart from the dynamical models mentioned above, Nagakawa & Nezu [34] and Perry & Chong [32] provide two of the very few kinematic models. The latter, in particular, has been refined over the years, and the latest in that series is given by Marušić *et al.* [35]. Also, Meng (chapter 10 in [1]) proposes a statistical approach which may prove

useful. The quantitative modeling of boundary layer structure seems a very fertile area for future work: we need to translate our descriptions on boundary layer behavior into quantitative models in order to claim that we understand boundary layer structure.

Finally, it is clear that one of the major difficulties with the data base is that laboratory studies are confined to a relatively small Reynolds number range. Detailed studies of boundary layer structure are limited, by and large, to momentum thickness Reynolds numbers less than about 15,000, whereas applications to vehicles requires extrapolation over several orders of magnitude beyond this value. More complete structural information in high Reynolds number flows will certainly help to clarify some of the outstanding scaling issues. The investigation of rough wall boundary layers can have a similar salutary effect on our understanding (see, in particular, the observations made by Antonia & Djenidi in chapter 9). The study of of non-canonical boundary layers (with roughness, pressure gradient, wall curvature, etc.) will undoubtedly test the robustness of our understanding, and help to advance the predictive character of our models.

Acknowledgments

This paper was prepared while AJS received funding from AFOSR Grant F49620-00-1-0139, NSF Grant CTS-9706902, and ONR Grants N0014-97-1-0618, N0014-98-1-0325, and N0014-99-1-0340.

References

1. Panton, R.L. (ed.) *Self-Sustaining Mechanisms of Wall Turbulence* Computational Mechanics Publications, Southampton, UK and Boston, USA, 1997.
2. Zagarola, M.V. & Smits, A.J. Scaling of the mean velocity profile for turbulent pipe flow. *Physics Review Letters*, **78** (2), 239–242, 1997.
3. Zagarola, M.V. & Smits, A.J. A new mean velocity scaling for turbulent boundary layers. *ASME Paper FEDSM98-4950*, 1998.
4. Robinson, S.K. The kinematics of turbulent boundary layer structure. *NASA TM 103859*, 1991.
5. Kline, S.J. & Robinson, S.K. Quasi-coherent structures in the turbulent boundary layer. Part 1. Status report on a community-wide summary of the data. *In:* Kline, S.J. & Afgan, N.H. (eds), *Near-Wall Turbulence.* Hemisphere, 1989.
6. Cantwell, B.J. Organized motion in turbulent flow. *Annual Review of Fluid Mechanics*, **13**, 457–515, 1981.
7. Smith, C.R., Walker, J.D.A., Haidara, A.H. & Sobrun, U. On the dynamics of near-wall turbulence. *In:* Walker, J.D.A. (ed), *Turbulent Flow Structure Near Walls.* The Royal Society. First published in *Phil. Trans. R. Soc. London A*. 336, 1991.
8. Morrison, J.F., Subramanian, C.S. & Bradshaw, P. Bursts and the law of the wall in turbulent boundary layers. *Journal of Fluid Mechanics*, **241**, 1992.
9. Smith, C.R. & Metzler, S.P. The characteristics of low-speed streaks in the near-wall region of a turbulent boundary layer. *Journal of Fluid Mechanics*, **129**, 27–54, 1983.

10. Spina, E.F., Donovan, J.F. & Smits, A.J. On the structure of high-Reynolds-number supersonic turbulent boundary layers. *Journal of Fluid Mechanics*, **222**, 293–327, 1991.
11. Smits, A.J. & Dussauge, J.P. *Compressible Turbulent Shear Layers.* Springer Verlag, AIP Imprint, 1997.
12. Morrison, J.F., Jiang, W-M., McKeon, B.J. and Smits, A.J. Scalings of the streamwise velocity components in turbulent pipe flow. In preparation, 2000.
13. Head, M.R. & Bandyopadhyay, P. New aspects of turbulent boundary layer structure. *Journal of Fluid Mechanics*, **107**, 1981.
14. Fernholz, H.H. & Finley, P.J. Incompressible zero-pressure-gradient turbulent boundary layers: An assessment of the data. *Progress in Aerospace Science*, **32**, 1996.
15. Smith, R.W. *Effect of Reynolds number on the structure of turbulent boundary layers.* Ph.D. Thesis, Princeton University, Princeton, NJ, 1994.
16. Brown, G.L. & Thomas, A.S.W. Large structure in a turbulent boundary layer. *Physics of Fluids*, **20** (10) Pt. II, 1977.
17. Falco, R.E. Coherent motions in the outer region of turbulent boundary layers. *Physics of Fluids*, **20** (10) Pt. II, 1977.
18. Smith, R.W. & Smits, A.J. Effect of Reynolds number on the large structure of turbulent boundary layers. *AIAA Paper 91-0526*, 1991.
19. Murlis, J., Tsai, H.M. & Bradshaw, P. The structure of turbulent boundary layers at low Reynolds numbers. *Journal of Fluid Mechanics*, **122**, 13–56, 1982.
20. Smits, A.J., Spina, E.F., Alving, A.E., Smith., R.W., Fernando, E.M. & Donovan, J.F. A comparison of the turbulence structure of subsonic and supersonic boundary layers. *Physics of Fluids A*, **1**, 1865–1875, 1989.
21. Luchik, T.S. & Tiederman, W.G. Timescale and structure of ejections and bursts in turbulent channel flows. *Journal of Fluid Mechanics*, **174**, 1987.
22. Bogard, D.G. & Tiederman, W.G. Burst detection with single-point velocity measurements. *Journal of Fluid Mechanics*, **162**, 1986.
23. Delo, C.J. *Volumetric Analysis of a Low Reynolds Number Turbulent Boundary Layer.* Ph.D. thesis, Princeton University, Princeton NJ, 1996.
24. Delo, C.J. & Smits, A.J.. Volumetric visualization of coherent structure in a low Reynolds number turbulent boundary layer. *International Journal of Fluid Dynamics*, **1**, 1997.
25. Goldstein, J.E. & Smits, A.J. Flow visualization of the three-dimensional, time-evolving structure of a turbulent boundary layer. *Physics of Fluids*, **6** (2), 1994.
26. Perry, A.E., Lim, T.T. & Teh, E.W.. A visual study of turbulent spots. *Journal of Fluid Mechanics*, **104**, 1981.
27. Acarlar, M.S. & Smith, C.R. A study of hairpin vortices in a laminar boundary layer. Part 1. Hairpin vortices generated by hemisphere protuberances. *Journal of Fluid Mechanics*, **175**, 1–41, 1987.
28. Acarlar, M.S. & Smith, C.R. A study of hairpin vortices in a laminar boundary layer. Part 1. Hairpin vortices generated by fluid injection. *Journal of Fluid Mechanics*, **175**, 43–83, 1987.
29. MacAulay, P. & Gartshore, I.P. A tentative model of outer-region structure in a turbulent boundary layer developing on a smooth wall. In *Experimental Heat Transfer, Fluid Mechanics and Thermodynamics 1991*, (J. F. Keller, R. K. Shah and E. N. Ganic, eds.), Elsevier, 1991.
30. Adrian, R. J., Christensen, K. T., Soloff, S. M., Meinhart, C. D. & Liu, Z.-C. Decomposition of turbulent fields and visualization of vortices. 9th Int'l Symp. Applications of Laser Techniques to Fluid Mechanics, Lisbon, July 13-16, 1998.

31. Robinson, S.K. Coherent motions in the turbulent boundary layer. *Annual Review of Fluid Mechanics*, **23**, 601–639, 1991.
32. Perry, A.E. & Chong, M.S. On the mechanism of wall turbulence. *Journal of Fluid Mechanics*, **119**, 173–219, 1982.
33. Aubry, N., Holmes, P. Lumley, J.L. & Stone, E. The dynamics of coherent structures in the wall region of a turbulent boundary layer. *Journal of Fluid Mechanics*, **192**, 115, 1988.
34. Nagakawa, H. & Nezu, I. On a new eddy model in turbulent shear flow. *Proc. of J.S.C.E.*, **231**, 61–70, 1974.
35. Marušić, I., Uddin, A.K.M. & Perry, A.E. Similarity law for the streamwise turbulence intensity in zero-pressure-gradient turbulent boundary layers. *Physics of Fluids*, **9** (12), 3718–3726, 1997.
36. Cantwell, B.J., Coles, D. & Dimotakis, P. Structure and entrainment in the plane of symmetry of a turbulent spot. *Journal of Fluid Mechanics*, **87**, 1978.
37. Falco, R.E. A coherent structure model of the turbulent boundary layer and its ability to predict Reynolds number dependence. *In:* Walker, J.D.A. (ed), *Turbulent Flow Structure Near Walls*. The Royal Society. First published in *Phil. Trans. R. Soc. London A*. 336, 1991.

The Largest Scales in Turbulent Flow: The Structures of the Wall Layer

Javier Jiménez[1,2] and Mark Simens[1]

[1] School of Aeronautics, 28040 Madrid, Spain
[2] Centre for Turbulence Research
Stanford University, Stanford CA, 94305–3030 USA

1 Introduction

Walls are present in most flows, such as in pipes, rivers, vehicles, etc., and they profoundly influence turbulence. Wall-bounded turbulent flows are of prime technological importance, but they are also physically very interesting. We shall see that many issues in them are still not understood, and that they differ in many aspects from isotropic and homogeneous flows. They are perhaps the last area in 'classical' incompressible turbulence in which there are still open questions about basic physical mechanisms.

We will focus in this paper on the structural organization of the wall region, and specially into the genesis of the structures in the viscous and logarithmic layers. A brief review of the classical theory is given first, for the benefit of those not familiar with the subject [37,40].

To a first approximation we can assume that the flow is homogeneous in the streamwise (x) and spanwise (z) directions. In the wall-normal (y) direction the wall makes the flow inhomogeneous. It also makes it anisotropic everywhere. This is due to two effects. In the first place the impermeability condition at the wall prevents wall-normal velocity fluctuations from developing length scales much larger than y, inducing a natural scale stratification in which larger structures are only present away from the wall. Also, the wall enforces a no-slip condition for the other two velocity components, so that viscosity cannot be neglected even at high Reynolds numbers.

The flow is for these reasons governed by two different length scales. Near the wall there is a thin layer in which viscosity dominates, but in which velocities are too small compared to the free-stream velocity, U_0, for the latter to be relevant. The velocity scale is there determined by the skin friction τ as

$$u_\tau^2 = \tau/\rho = \nu \partial_y U, \tag{1}$$

where ρ is the fluid density, ν is the kinematic viscosity, and U is the mean velocity profile. The length scale is given by ν and u_τ, and variables normalized in these 'wall' units are traditionally distinguished by a $^+$ superindex. The dimensionless wall distance is, for example, $y^+ = u_\tau y/\nu$, and the dimensionless flow thickness $Re_\tau = u_\tau h/\nu$ is a good characterization of the Reynolds number of the flow. Typical values are $Re_\tau \approx 100$ for transitional flow to $Re_\tau \approx 10^4$ for

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 39–57, 2001.
© Springer-Verlag Berlin Heidelberg 2001

geophysical ones. In this near-wall region the velocity is given by

$$U^+ = f(y^+), \tag{2}$$

where f is expected to depend only on wall properties (smoothness, etc.)

Away from the wall, where $y = O(h)$, the Reynolds number is too high for viscosity to be important. The length scale there is the flow thickness (e.g. the pipe radius), but the velocity scale for the velocity fluctuations is still u_τ. This is because momentum conservation implies that τ is transmitted more or less unchanged across the layer by the Reynolds stresses, $-\langle u'v' \rangle$. The velocity has been separated here into a mean value and fluctuations

$$\mathbf{u} = [U(y),\, 0,\, 0] + [u',\, v',\, w'], \tag{3}$$

and $\langle\rangle$ stands for temporal (or for ensemble) averaging. The streamwise, wall-normal and spanwise velocity components are u', v' and w'. The magnitude of the Reynolds stresses suggests that $u' = O(u_\tau)$ and that the velocity differences in this region are also of the same order. The general form of the velocity profile in this layer is then

$$U_0 - U = u_\tau F(y/h), \tag{4}$$

An overlap layer is needed to match the length scales in the inner and outer flows, whose ratio is $Re_\tau \gg 1$. In this layer the distance to the wall is too small for h to be relevant, $y/h \ll 1$, while the Reynolds number is too large for the viscosity to be important, $y^+ \gg 1$. The only remaining length scale is y itself, and the velocity profile can be obtained from dimensional arguments. The only possible expression for the velocity gradient is

$$\partial_y U = \frac{u_\tau}{\kappa y}, \tag{5}$$

where the empirical Kármán constant, $\kappa \approx 0.4$, is in principle universal. Integration of this differential equation yields the classical logarithmic profile

$$U^+ = \frac{1}{\kappa} \log y^+ + A. \tag{6}$$

where the constant A has to be determined independently. This additive constant controls the overall friction coefficient since, if we assume that (6) holds up to the centre of the pipe, $y = h$, we obtain u_τ in terms of A and of Re_τ. The viscous layer lies underneath the level in which (5) applies, and acts as the boundary condition which fixes A, and the drag. The empirical limits for the logarithmic layer are $y/h < 0.2$ and $y^+ > 150$ [25]

Experiments agree very well with these scaling laws [41], even if the assumptions that we have made are only true to the lowest order, and this classical theory can be safely used as a framework for the study of the wall.

2 The Balance of Turbulent Energy

There is an alternative way of understanding the structure of wall flows. Assume a streamwise-uniform flow, such as a turbulent pipe, and separate the velocity into its mean value and fluctuations. We can write an equation for the evolution of the mean turbulent kinetic energy, $K = \langle u'^2 + v'^2 + w'^2 \rangle / 2$,

$$\partial_y \Phi_T = -\langle u'v' \rangle \partial_y U - \varepsilon + \nu \, \partial_{yy} K. \tag{7}$$

In this equation the first term in the right-hand side is the production of turbulent energy by the interaction between the mean shear and the Reynolds stresses, the last one is the large-scale viscous diffusion of the kinetic energy, which can be neglected everywhere except in the viscous sublayer, and ε is the dissipation by the small scales $\nu \langle |\nabla \mathbf{u}|^2 | \rangle$. The balance of these three terms is compensated in the left-hand side by the divergence of a spatial energy flux

$$\Phi_T = \frac{1}{2} \left\langle v' (u'^2 + v'^2 + w'^2) \right\rangle + \langle v' p' \rangle, \tag{8}$$

where p' are the pressure fluctuations.

This equation provides an alternative derivation of the logarithmic velocity profile [40, p. 135]. If we assume that the energy production and dissipation are in equilibrium within the logarithmic region, and take the latter to be $u_\tau^3 / \kappa y$, the energy balance becomes,

$$u_\tau^2 \partial_y U = u_\tau^3 / \kappa y, \tag{9}$$

which integrates directly to (6). The left-half side of this equation is an exact expression for the energy production, while the right-hand side depends only on generic properties of the turbulent cascade. The assumption is that the cascade is implemented by approximately isotropic eddies whose velocities are $O(u_\tau)$ and whose sizes are $O(y)$. Note that this assumption is equivalent to the length scaling hypothesis used to derive (6), but that here it gives a more concrete interpretation of the Kármán constant κ as defining the integral scale of the eddies of the logarithmic region.

A consequence of this analysis is that the energy flux Φ_T should be constant across the logarithmic layer, and this is tested in figure 1.a, which contains data from three numerical channels at moderate Reynolds numbers. Positives fluxes flow away from the wall, and an increase of the flux with y implies a local excess of energy production. The three curves collapse well near the wall, where energy production dominates. Part of the excess energy diffuses into the wall, but the rest flows through the logarithmic layer into the outer flow, where dissipation is dominant. In this sense, the wall provides the power needed to maintain turbulence in the outer region. There is a plateau in the logarithmic region, in agreement with (9), together with perhaps a second energy production peak near the outer edge of the logarithmic layer.

It is unfortunately difficult to measure all the terms of the energy equation in experimental flows, and there are no data equivalent to those in figure 1**(a)** at

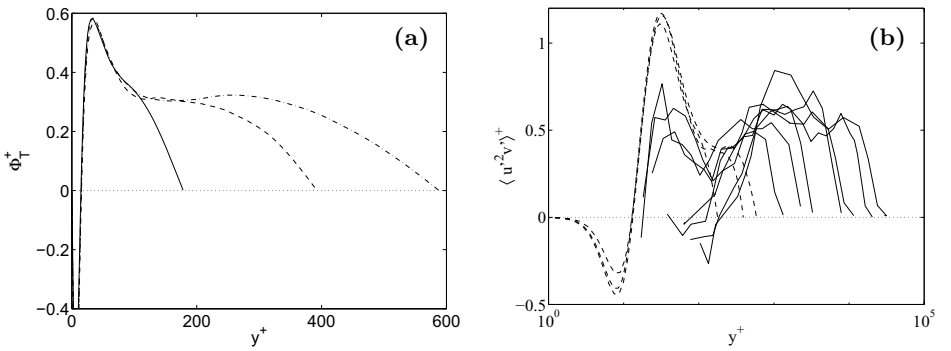

Fig. 1. Spatial flux of turbulent energy in wall bounded flows, normalized by u_τ^3. **(a)** Total flux computed from integrating equation (7) in numerical plane channels. ——— , $Re_\tau = 180$ [20]; – – – – , $Re_\tau = 395$; —·— , $Re_\tau = 590$ [27]. **(b)** Flux of streamwise turbulent fluctuations. The dashed lines are the three channels in **(a)**. The solid ones are boundary layers from [8]. $Re_\tau \approx 1,200 - 3 \times 10^4$.

higher Reynolds numbers. Some parts of Φ_T have however been measured and can be used to check how representative are the numerical results. This is done in figure 1**(b)**, which shows the turbulent transfer of streamwise fluctuations, $\langle u'^2 v' \rangle$, which accounts for about half of the flux in (8). That figure includes the numerical channels and experimental boundary layers at much higher Reynolds numbers. Notwithstanding the uncertainties in the experimental data, specially where the probe length is comparable to the wall distance, both set of measurements coincide where they overlap, giving some confidence on the conclusions from the numerics.

The picture suggested by this discussion is that of an energy producing region near the wall ($y^+ < 50$), which exports some of its extra turbulent energy across the logarithmic region into the outer flow. This spatial energy transport is an energy cascade different from the usual Kolmogorov one. In the latter, energy generated at some place is transferred locally to the smaller scales, where it is eventually dissipated by viscosity. In the spatial cascade the energy is not dissipated locally, but is transferred to other locations where it is eventually dissipated by cascading to smaller scales.

This second cascade is only found in inhomogeneous flows, and does not involve dissipation. In fact, since we have seen that in wall turbulence it flows away from the wall, where the scales are larger, it is actually a reverse cascade. Both the spatial and the Kolmogorov cascades are present in wall flows. It is difficult to compare their magnitude since Φ_T, which is an energy flux per unit area, has different units from ε, which is a flux per unit volume. We can however form a length from the two quantities, $\lambda_T = \Phi_T / \varepsilon$, which is the thickness of the layer over which the Kolmogorov cascade would be able to dissipate all the energy provided by the spatial flux. It follows from figure 1**(a)**, and from the estimates in (9), that $\lambda_T \approx 0.15y$, so that the turbulent energy being transferred

through each level of the logarithmic layer is equivalent to the energy being generated in the top 15% of its distance to the wall.

3 The Near-Wall Region

Because of this energy-generating property, and because we have seen that the magnitude of wall drag is controlled there, the region below $y^+ \approx 100$ has been the subject of intensive study. In this layer the wall-normal Reynolds numbers are low, and we may expect quasi-laminar structures whose behaviour can be understood deterministically. In some sense this region corresponds to the Kolmogorov viscous range of isotropic turbulence, and it is easy to show that the Kolmogorov and the wall-unit length scales are identical.

The dominant structures in this region are streamwise velocity streaks and quasi-streamwise vortices. The former were the first ones to be recognised [19], and consist of long ($x^+ \approx 1000$) sinuous arrays of alternating streamwise jets superimposed on the mean shear, with an average spanwise separation $z^+ \approx 100$ [33]. At the spanwise locations where the jets point forward, the wall shear is higher than the average, while the opposite is true for the 'low velocity' streaks where the jets point backwards. The net effect is to increase the mean wall shear, and it is the immediate cause of the turbulent wall drag [28]. Many attempts to control drag have therefore centred on weakening these near-wall structures [5,14].

The quasi-streamwise vortices are slightly tilted away from the wall, and each one stays in the near-wall region only for $x^+ \approx 200$ [13]. Several vortices are associated with each streak, with a longitudinal spacing of the order of $x^+ \approx 400$ [16]. Some of them are connected to the trailing legs of the vortex hairpins of the outer part of the boundary layer [29], but most merge into disorganised vorticity after leaving the immediate wall neighbourhood [32].

It has long been recognised that the vortices cause the streaks by advecting the mean velocity gradient [3]. This process is independent of the presence of the wall, as shown by the observation in [24] of streaks in uniformly sheared flows.

There is less agreement on the mechanism by which the vortices are generated. It was already proposed in [19] that streaks and vortices were part of a cycle, in which the latter were the result of an instability of the former. The argument was made more specific by [34], who noted that the spanwise velocity gradients that separate the low- from the high-velocity streaks are subject to inflectional instabilities. This conceptual model has been elaborated and used to explain several properties of disturbed boundary layers by [14,12] among others, and is the essence of the quasilinear models developed at MIT and Cornell, and reviewed in [22,10]. It can be summarised as that the quasi-streamwise vortices act on the mean shear to create the streaks, which become inviscidly unstable and eventually produce tilted streamwise vortices.

A slightly different point of view is that the cycle is organized around a nonlinear travelling wave, which would represent a permanently disturbed streak. This is not so different from the previous instability model, which essentially

assumes that the undisturbed streak is a fixed point, and that the cycle is a homoclinic orbit running through it. Candidate nonlinear waves have been computed by [36] and others, and identified as part of a particular path to turbulent transition in [38]. Finally, reduced models based on this approach have been formulated in [35]. A difficulty with most of these studies is that there is no clear connection between the object being identified and full-scale turbulence. What is needed is a demonstration that a turbulent flow can be continuously modified into one of these reduced objects while still remaining identifiably turbulent.

A possible way of doing this was proposed by [16], who substituted a full turbulent channel with an array of identical computational boxes, periodic in x and z, but each with its full wall-normal extent. The idea was to substitute the complexity of all the mutually interacting turbulent units near the wall by a 'crystal' of identical structures, all of them executing synchronously the turbulence regeneration cycle. The 'unit cell' of the crystal was adjusted to the smallest dimension that would maintain turbulence, which turned out to contain a single wavelength of a velocity streak and a pair of quasi-streamwise vortices. These structures went through a complex cycle which was still difficult to analyze, and the statistics of the fluctuations were essentially identical to those of fully developed channels. This 'minimal' system has often been used as a surrogate for real wall turbulence, and most of the reduced models mentioned in the last paragraph actually refer to it.

The previous experiment eliminated the possibility that wall turbulence required the interaction between neighbouring wall structures, but did nothing about the interaction with the core flow, which was still present. The question of the necessity of inner-outer interaction in wall turbulence has been a matter of much discussion. It is important in formulating scaling laws for the different regions, and in deciding which parts of the flow to target when designing turbulence control strategies.

An important step in clarifying this question was taken by [17], who eliminated the vorticity fluctuation in the outer flow of a numerical channel by using a numerical filter which acted as a variable viscosity, low near the wall and high outside. They were able to show that the wall cycle was 'autonomous' in the sense that it could run independently of the outer flow as long as it was not filtered below $y^+ \approx 60$. Again the statistics were similar to those of full turbulence.

The experiment of running a system that is both 'minimal', in the sense that its wall-parallel size is small enough to contain a single copy of the structures, and 'autonomous' in the sense of having no outer flow, requires some care, and had not been achieved until now. Some preliminary results are discussed next.

In all the numerical experiments the vertical filter was set to $y^+ = 50$, and the wall was run until it achieved statistically steady state. Tests were run with progressively smaller wall-parallel wavelengths, until enough flow simplification was achieved. This happened with for $x^+ \times z^+ = 500 \times 220$, which is somewhat longer and somewhat wider than the minimal unit. The flow entered a regular limit cycle, as opposed to the chaotic behaviour characteristic of full, or even of minimal full-depth channels. Decreasing the box side made the cycle weaker and

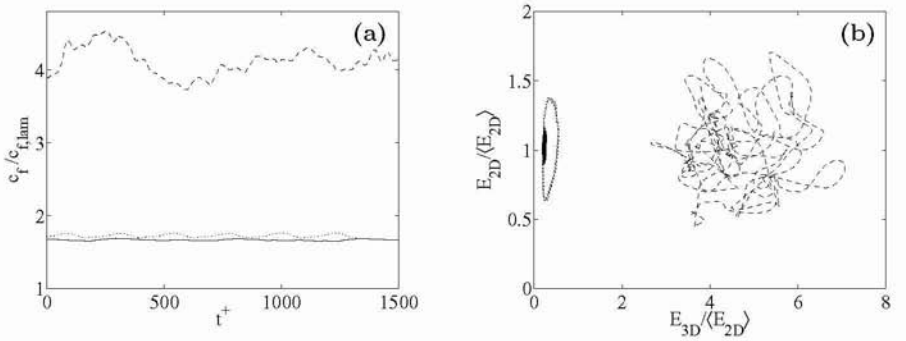

Fig. 2. Numerical experiments on minimal autonomous walls. – – – – , full-depth channel, $x^+ \times y^+ \times z^+ = 770 \times 380 \times 340$; ····· , $500 \times 50 \times 220$; ——— , $300 \times 50 \times 180$. All cases are shown after reaching statistically steady state. (**a**) Friction coefficients compared with the theoretical laminar value. (**b**) Perturbation energies defined in the text. There are seven cycles in the dotted line.

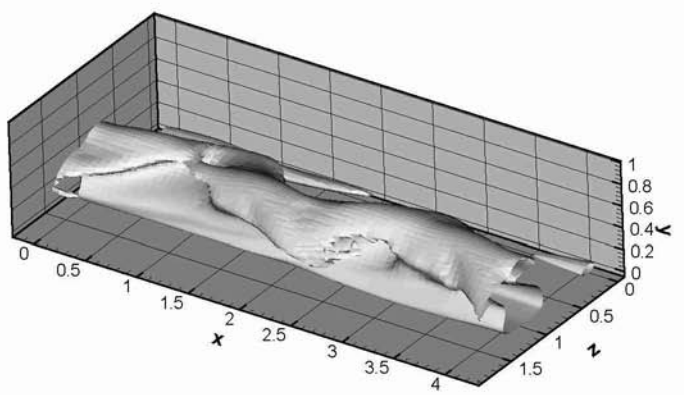

Fig. 3. Isosurface of the streamwise perturbation velocity for the intermediate box in figure 2. $u'^+ = -1$. Flow is from left to right.

drove it almost to a permanent travelling wave. The time evolution of the friction coefficients of two such boxes is shown in figure 2(**a**), which also contains data from a full-depth channel for comparison. Figure 2(**b**) shows the same data in a different representation. Orbits are shown in terms of the two-dimensional and three-dimensional energy of the v' fluctuations, as first used in [38]. The former is the energy of the fluctuations of the streamwise-averaged flow, and therefore represents the intensity of long objects, essentially streaks. The three-dimensional component contains the deviations from streamwise uniformity. The figure underscores the difference between the limit cycle and the chaotic evolution of the full channel, and shows that the simplification process in constrained systems corresponds to a decrease of the three-dimensionality of the flow. In agreement with this, and with previous results in [16,17], the fluctuation intensity profiles

for these flows are essentially normal for the streamwise velocity component, but weaker than normal for the other two components.

A three-dimensional representation of the streak in the longer of the two autonomous boxes of the previous figure is shown in figure 3. It shows a streak with two waves, associated to each of which there is a pair of quasi-streamwise vortices (not shown). During the turbulent cycle these waves grow deeper and shallower, but the streak is never uniformly straight. The presence of two waves suggests that it should be possible to have an autonomous flow with a single one, and the shorter box in figure 2 is indeed one such solution. Even if the present flows are not permanent waves, they are very weakly oscillating ones, and it should be emphasized that they are statitistically stationary and autonomous, in the sense that there are no other vorticity fluctuations in the numerical box except those seen in the figure. They have also been derived from a fully turbulent flow by a continuous set of transformations in which their statistical properties were either maintained or evolved smoothly. They are the best representation that we have up to now of the elemental 'engine' which powers near-wall turbulence.

Figure 3 is qualitatively very similar to the traveling wave obtained in [36], and to the structures educed from larger channels by [16,17]. Casual observation of the differences between these and larger flows suggest the root of the complexity in the latter. As the streak goes through the cycle, it ejects some vorticity into the outer flow, essentially as a small hairpin vortex. In full-depth flows this vorticity evolves, becomes disorganized, and eventually comes back to modify the next cycle of the streak. In autonomous flows the vorticity is damped by the filter as soon as it is ejected, and this randomizing mechanism is not present.

4 The Structures of the Overlap Layer

Although the structures of the near-wall region have been the subject of a lot of work, the overlap layer has been studied mostly from the statistical point of view. The reason has been in part that it is difficult to obtain experimentally flows with a clean overlap region. If we follow the criterion that the latter exists only above $y^+ \approx 100 - 200$ and below $y/\delta \approx 0.15 - 0.2$, the minimum Reynolds number for which some overlap can be expected is $Re_\tau = 500$, and at least $Re_\tau = 1000$ is needed to have some confidence in the results. Most structural studies, which typically require visualizations and measurements over extended regions, are made at lower Reynolds numbers. This is also true of numerical simulations. Moreover, even if the Reynolds number is high enough, the overlap structures are large and themselves turbulent. This makes them harder to visualize, and even to define, than the smoother ones near the wall.

Most of the information on this region comes from velocity spectra, and suggests the presence of extremely large structures [4,7,30]. A compilation of premultiplied spectra of the streamwise velocity fluctuations, $k_x E_{uu}(k_x)$, is given in figure 4**(a)**, and they extend to wavelengths, $\lambda = 2\pi/k$, of the order of $60 - 100y$. At the edge of the overlap region, this corresponds to $15 - 20$ times the boundary layer thickness. This long-wave k_x^{-1} energy spectrum of wall flows has

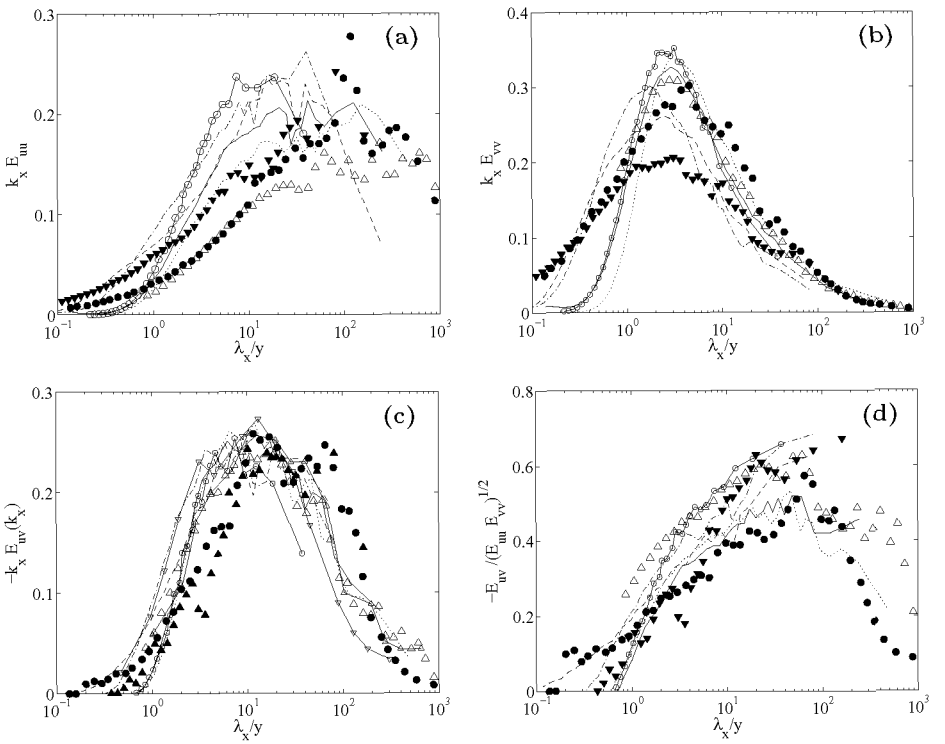

Fig. 4. Premultiplied velocity spectra, as a function of streamwise wavelength $\lambda_x = 2\pi/k_x$, normalized to integrate to unity. Abscissae are normalized with wall distance. **(a)** E_{uu}. **(b)** E_{vv}. **(c)** $-E_{uv}$. **(d)** Structure cross-correlation parameter $-E_{uv}/(E_{uu}E_{vv})^{1/2}$. —○— , numerical channel, $u_\tau\delta/\nu = 590$, $y^+ = 100$ [27]. —△— , $y^+ = 100$; —▽— , $y^+ = 250$, pipe with $u_\tau R/\nu = 2100$ [23]. Simple lines are laboratory smooth-wall boundary layers from Carlier (private commun.): ····· , $y^+ = 95$; - - - - , $y^+ = 360$. $u_\tau\delta/\nu = 2530$. ——— , $y^+ = 110$; —·— , $y^+ = 970$. $u_\tau\delta/\nu = 7050$; Symbols are from the atmospheric boundary layers: △ , $y/h = 100$; • , $y/h = 1000$; from [11], $U \approx 10$ m/s at 10 m, and roughness height, $u_\tau h/\nu \approx 400$. If we assume the boundary layer thickness to be $\delta \approx 500$ m, $\delta^+ = \delta/h \approx 5 \times 10^4$. ▼ , from [1] $y = 1.7$ m, $U = 6.2$ m/s, and roughness height, $h = 4$ mm.

been extensively studied in [29], and is only present at wall distances which more or less coincide with the overlap layer. Since the total energy of the fluctuations can be written as

$$\langle u^2 \rangle = \int_0^\infty k_x E_{uu}(k_x)\, \mathrm{d}\log\lambda_x, \tag{10}$$

premultiplied spectra extending to long wavelengths imply that a large fraction of the total fluctuating kinetic energy is associated with those structures.

Even if most of this information has been available for some time, it was not until the compilations presented in [15,21], or until the publication of the thesis [9] that the real size of these structures became widely realized. A consequence

has been that many experimental measurements, and most computations, were not planned with these structures in mind, and that the records compiled in many of them are too short, giving the false impression that the structures are shorter than they really are.

A compilation of spectra for the wall-normal velocity component is also given in figure 4, which emphasizes the paucity of available data for E_{vv} and for the E_{uv} cospectrum. The latter is specially important, since its integral is the transverse Reynolds stress $-\langle u'v' \rangle$, which in turn determines the mean velocity profile. Only if this quantity scales well with the distance to the wall is there any reason to expect the self-similar scaling of the Reynolds stresses which is needed for the logarithmic law. It is seen in figures 4 that the E_{vv} and E_{uv} spectra actually scale better with wall distance than E_{uu}. This is especially true for the cospectrum, even if the experiments used in the figure are very different from each other, giving strong support to the logarithmic velocity profile. Older compilations of stress cospectra for more homogeneous conditions can be found in [26,2].

It is interesting that the wavelengths implied by the spectra for the Reynolds stresses are longer than those the wall-normal velocity, even if the former, $-\langle u'v' \rangle$, cannot exist without the latter. The reason is clarified by the spectral distribution of the structure parameter in figure 4**(d)**, which describes the *efficiency* of the turbulent fluctuations in transporting momentum. If this quantity is low, the u and v fluctuations are uncorrelated, and only produce low stresses, while the opposite is true if it is high. The distribution in figure 4**(d)** is consistent with the intuitive idea that isotropic turbulence is unable to transmit momentum. For wavelengths $\lambda_x \approx y$ the efficiency parameter is almost zero, since these scales are small enough not to be influenced by the wall, and are therefore almost isotropic. It is only for the longer structures, which feel the wall and the mean shear, that the efficiency approaches unity and that the Reynolds stresses become significant. It is this bias of the efficiency towards longer wavelengths that shifts the peak of the cospectrum with respect to that of E_{vv}.

Much less is known of the spanwise extent of these fluctuations. The spectra in figure 5 suggest that both u and v have lateral wavelengths which are low multiples of the distance to the wall, but they were compiled from numerical simulations at low Reynolds numbers where a proper logarithmic layer does not exist. They also correspond to computational boxes that, as we have noted above, are too short to contain the longest streamwise wavelengths, and it is possible that those long waves may also be wide, in which case they would be missed by the numerical spectra. This is suggested by the two-dimensional spectra in figure 6, which correspond to the same numerical data, and which show a 'diagonal' trend for both E_{uu} and E_{ww}, suggesting that longer and wider structures could be present in larger numerical boxes. To our knowledge there are no published experimental spanwise spectra on which to check this possibility, but there is at least one report, from the analysis of PIV data in a turbulent channel, which claims that most of the energy in E_{uu} very close to the wall is in very wide spanwise wavelengths [39].

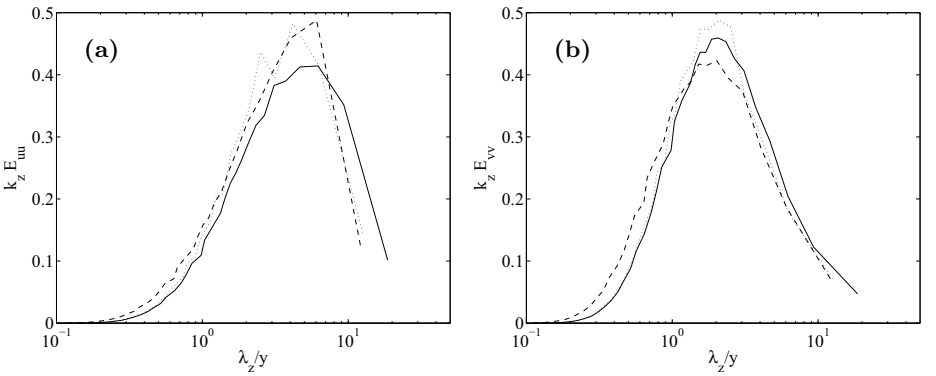

Fig. 5. Premultiplied velocity spectra, as a function of spanwise wavelength $\lambda_z = 2\pi/k_z$, normalized to integrate to unity. Abscissae are normalized with wall distance. Numerical channel with $Re_\tau = 590$, ——— , $y^+ = 100$; - - - - , $y^+ = 150$; ····· , $y^+ = 100$ at $u_\tau\delta/\nu = 395$ [27]. **(a)** E_{uu}. **(b)** E_{vv}.

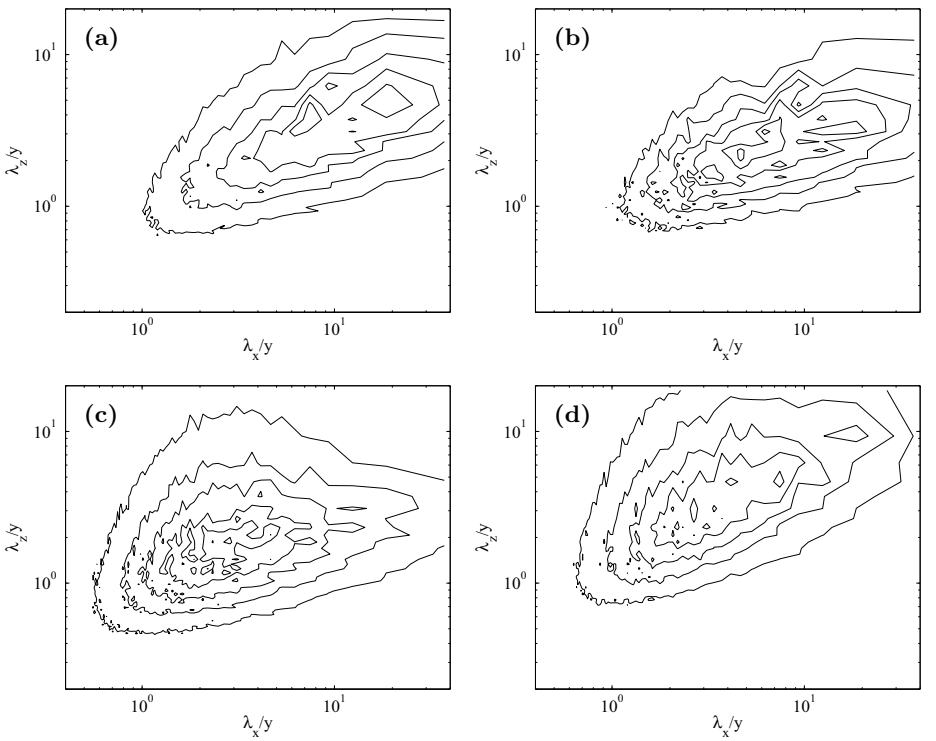

Fig. 6. Premultiplied two-dimensional velocity spectra, normalized to integrate to unity, as a function of the streamwise and spanwise wavelengths normalized with wall distance. Numerical channel from [27]. $Re_\tau = 590$, $y^+ = 100$. **(a)** E_{uu}. **(b)** $-E_{uv}$. **(c)** E_{vv}. **(d)** E_{ww}. Contour intervals are 0.01 (0.01) for E_{uu}, and 0.03 (0.03) for the other three spectra.

From the evidence of the spectra in this section, at least until the issue of the spanwise wavelengths is clarified, the geometry of the large structures of the overlap layer is that of long and narrow regions which exist only for the streamwise, but not for the wall-normal, velocity component, and which can therefore be described as streaks, broadly similar to those of the near-wall region. They are however much larger than the latter, and fully turbulent rather than laminar or transitional. Some references to direct observations of these large streaky structures, in nature and in numerical simulations, can be found in [15].

5 Linear Stability of the Mean Velocity Profile

The size of the structures discussed in the previous section suggests that they should be describable in terms of linear processes. Since we have observed that they contain much of the kinetic energy, which is of the order of u_τ^2, their characteristic velocity scale is u_τ, and their turnover time is $T_L = L/u_\tau$, where L is their characteristic length. This self-deformation time competes with the shearing time scale of the mean velocity profile which, from the logarithmic law, is $T_s = (\partial_y U)^{-1} = \kappa y/u_\tau$. For any structure for which $L \gg y$, including most of those discussed in the previous section, T_L is much slower than T_s, and the nonlinear self-interactions are not relevant for the dynamics.

The traditional emphasis when analyzing linear structures has been on unstable eigenmodes, since it is clear that, if they exist, they will grow and be observed. These are commonly associated with inflection points of the mean profile [6], in which case their growth rates are of the order of the local shear, and they compete with other processes present in the turbulent flow.

The mean velocity profiles of most attached wall flows have no inflection points, and are believed to be stable. This was shown for example in [31] for a turbulent channel with impermeable walls. They used a model velocity profile which approximated those of experimental channels, and substituted the molecular viscosity by the isotropic eddy viscosity which was needed to maintain the profile. Both approximations need some comment. The first one assumes that something like a mean profile exists instantaneously, which is known not to be the case. The second one assumes that the Reynolds stresses are carried by structures which are much smaller than those being studied. In both cases the implicit assumption is that there is a separation of scales between the large structures and the 'small scale' turbulence.

Such spectral gaps do not generally exist, since turbulent spectra are broadbanded, and such assumptions have to be justified in each particular case. There are two reasons why we will use it here. In the first place, and as long as we restrict ourselves to analyzing the very long u structures, the spectra in the previous section suggest that the Reynolds-stress producing scales are indeed shorter by an order of magnitude, allowing us to plausibly model them by an isotropic eddy viscosity. Note that this argument would not be true for the v structures, whose spectral peaks are at shorter wavelengths than those of E_{uv}. This suggest that we should use the mean-field, eddy-viscosity approximation,

estimate from it the Reynolds stresses, and use only results for those wavelengths which are clearly longer than the stress-producing ones.

The second reason for using this approximation is that we shall see that the eigenfunctions of interest are essentially neutral inviscid modes, and that the role of the eddy viscosity is to damp them slightly. We may therefore expect the eigenfunctions to be essentially correct even if the magnitude of the imaginary damping part of the eigenvalue is not accurate.

With these limitations in mind we will next analyze the stability of wall flows using linear perturbations of the mean profile, and substituting the constant molecular viscosity by the variable isotropic eddy viscosity,

$$\nu_t = u_\tau^2 \, \frac{1 - y/h}{\partial_y U}. \tag{11}$$

Note that this definition of ν_t includes the molecular viscosity. Assume perturbation velocities of the form

$$v' = \widehat{v}(y) \exp \mathrm{i}[\alpha_1 (x - ct) + \alpha_3 z], \qquad \text{etc.}, \tag{12}$$

where the wavenumbers α_1 and α_3 are real, and c is an unknown phase velocity, possibly complex. The linearized equation takes the form

$$(U - c)\,(\partial_{yy} - \overline{\alpha}^2)\,\widehat{v} - U_{yy}\widehat{v} + \mathrm{i}\alpha_1^{-1}(\partial_{yy} - \overline{\alpha}^2)[\nu_t(\partial_{yy} - \overline{\alpha}^2)\,\widehat{v}] = -\mathrm{i}\alpha_1 \nu_{t,yy}\widehat{v}, \tag{13}$$

where $\overline{\alpha}^2 = \alpha_1^2 + \alpha_3^2$, and the subindices of the mean velocity and eddy viscosity denote derivatives with respect to y. The left-hand side is the standard Orr-Sommerfeld stability equation [6], while the extra term in the right-hand side originates from the variation of the eddy viscosity, and can be shown numerically to affect the solutions only slightly. The no-slip boundary condition can be written, using continuity, as

$$\partial_y \widehat{v}(2h) = \partial_y \widehat{v}(0) = 0, \tag{14}$$

while the impermeability condition is

$$\widehat{v}(2h) = \widehat{v}(0) = 0. \tag{15}$$

We will substitute it here, for reasons that will later become clear, by the more general form,

$$\widehat{v} + \beta_0 \widehat{v} + \beta_1 \partial_y \widehat{v} = 0, \tag{16}$$

where β_0 and β_1 are arbitrary small parameters.

Except for the right-hand side of (13), Squire's transformation [6] reduces the problem (13)–(15) to a two-dimensional one with $\alpha_3 = 0$, and a modified viscosity

$$\widehat{\nu}_t = \overline{\alpha} \nu_t / \alpha_1, \tag{17}$$

Oblique waves therefore behave as normal ones with the same $\overline{\alpha}$ and higher viscosity and, since the latter is generally a stabilizing effect, we may assume

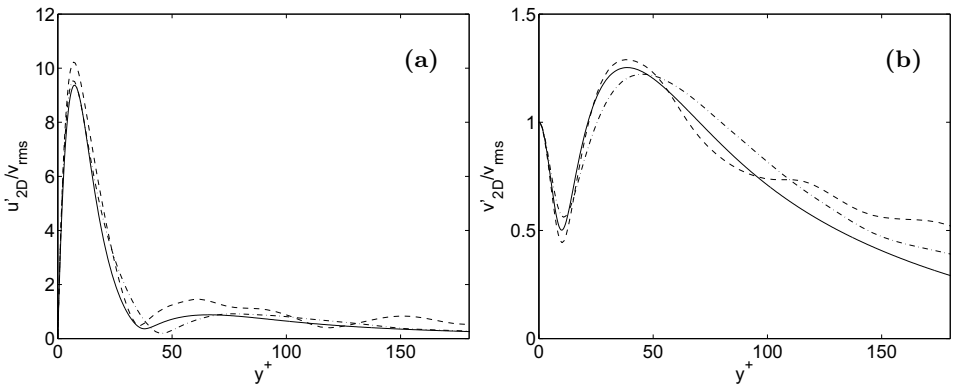

Fig. 7. Profiles of the spanwise-coherent fluctuation intensities, normalized with the forcing v_{rms} at the wall. Phase velocity of the forcing $U_c^+ = 10$. Wavelength $\lambda_x/h = 2.7$. $h^+ \approx 200$. - - - - , $v_{rms}^+ = 0.1$; —·— , $v_{rms}^+ = 0.2$; ——— , linear calculation. **(a)** $\langle u^2 \rangle^{1/2}$. **(b)** $\langle v^2 \rangle^{1/2}$.

for the moment that $\alpha_3 = 0$. The right-hand side of (13) does not transform correctly under (17), so that the transformation is not exact, but we have seen that its effect is small.

The accuracy of the eddy-viscosity approximation can be tested by forcing experiments in which a sine-wave transpiration is artificially imposed in one of the walls. This is equivalent to substituting (15) by $\widehat{v}(0) = 1$. The result, compared with direct simulations of forced turbulent channels, is shown in figure 7, which is adapted from [18]. The linear scaling of the velocity perturbation profiles is excellent for two different forcing intensities, and so is the agreement with linear theory, even if the forcing in the simulations was strong enough to change the magnitude of the skin friction by almost 40%. Other quantities are however not well predicted. The increment of the Reynolds stresses due to the forcing is, for example, underpredicted by a factor of about two.

5.1 A Simple Inviscid Profile

The general properties of the stability problem can best be understood by the inviscid analysis of a piecewise-linear velocity profile over a single wall,

$$U = U_0 y/h \quad \text{if} \quad y < h; \qquad U = U_0 \quad \text{otherwise.} \tag{18}$$

In the inviscid case (13) reduces to a second-order equation, and the only boundary conditions are the generalized impermeability condition (16) at the wall, and $\widehat{v} \to 0$ as $y \to \infty$. The no-slip condition is lost. The solution is continuous everywhere and can be expressed as a combination of exponentials $\exp(\pm \overline{a} y)$. The discontinuity of $\partial_y \widehat{v}$ at $y = h$ is computed by integrating (13) across the corner of the profile. There is a single eigenvalue which can be expressed as a function of the wavenumber magnitude, $\overline{\alpha}$, and of the permeabilities.

In the impermeable case, $\beta_0 = \beta_1 = 0$, the eigenvalue is neutral. The phase velocity is real and equal to

$$c_r/U_0 = \left(\mathrm{e}^{-2\overline{\alpha}h} + 2\overline{\alpha}h - 1\right)/2\overline{\alpha}h. \tag{19}$$

For very long waves $c_r/U_0 \approx \overline{\alpha}h \ll 1$, and the corresponding eigenfunctions decay only slowly at large y. For short waves, $\overline{\alpha}h \gg 1$, the phase velocity is $c_r \approx U$, and the eigenfunctions are localized in the neighbourhood of the corner of the profile. This mode corresponds to a neutral shear wave, and exists for all boundary layer profiles in the inviscid limit. Viscosity weakly damps it, but it is always there, potentially ready to couple with the permeability boundary condition (16).

For $\beta_i \ll 1$ we can indeed expand the eigenvalues as a series of the boundary condition coefficients. The lowest order correction is linear in β

$$\Delta c/U_0 = \beta_1 \mathrm{e}^{-2\overline{\alpha}h}/h. \tag{20}$$

The change in the real part of the phase velocity is $O(\beta)$ and can be neglected to the lowest order, but the stability is determined by the imaginary part of β_1. The unstable eigenfunctions are also $O(\beta)$ perturbations of the neutral mode of the homogeneous boundary conditions.

5.2 An Example: The Porous Wall

A direct application of the previous stability analysis was found in the numerical experiments of turbulence over porous walls carried out in [18]. The porosity of the wall is modelled by a Darcy-type law

$$v' = -\gamma p', \tag{21}$$

where γ is a porosity coefficient with dimensions of an inverse velocity. It is zero for impermeable walls, and tends to infinity for completely permeable ones.

The pressure at the wall is given by

$$\overline{\alpha}^2 \widehat{p} = \partial_y[\nu_t(\partial_{yy} - \overline{\alpha}^2)\widehat{v}] + \mathrm{i}\alpha_1 \left(U_y \widehat{v} + c\partial_y \widehat{v}\right). \tag{22}$$

In the inviscid case (21) has therefore the general form (16) with $\beta_1 = \mathrm{i}\gamma\alpha_1 c_r/\overline{\alpha}^2$, where c_r can, to lowest order, be substituted by (19). Since the imaginary part of β_1 is positive, its effect is destabilizing and we expect large-scale instabilities, which were observed. The predicted growth rate for low porosities is

$$\alpha_1 \Im(c) h/U_0 = \gamma \alpha_1^2/\overline{\alpha}^2 c_r. \tag{23}$$

The ratio of the wavenumbers is maximum when $\alpha_3 = 0$, and we therefore expect to observe, as we do, spanwise oriented perturbations.

There is in this case a simple interpretation for the observed instability. When $\gamma U_0 \gg 1$ the porous boundary condition (21) is equivalent to $\widehat{p}(0) = 0$ which, using (22), reduces to $\partial_y \widehat{v}(0) = 0$. If we extend antisymmetrically the profile

(18) to $y < 0$, it becomes a piecewise-linear free shear layer, whose stability characteristics are well known (see [6, p. 146]). The boundary condition at $y = 0$ fixes the parity of the solution. The $\widehat{v}$ eigenfunctions of the infinite-porosity case are even with respect to $y = 0$, inducing sinuous deformations of the shear layer. These are the well-known Kelvin-Helmholtz instability waves, and they are the only instabilities of this profile. The varicose deformation generated by the impermeable boundary condition, $\widehat{v}(0) = 0$, is neutral. The intermediate porosities connect the Kelvin-Helmholtz instability of the fully permeable wall to the varicose neutral modes of the impermeable one.

A word of caution is needed here. The previous inviscid analysis of a model boundary layer can be repeated for a channel [18], and the results are somewhat different. There are in that case two initially neutral modes, which agree for short wavelengths, but which differ substantially for long ones, suggesting that the stability characteristics of internal and external flows might not be exactly equal in that limit.

5.3 Viscous Modes

The important conclusion that can be drawn from the previous discussion is that there are weakly damped eigenfunctions of the mean velocity profile of wall-bounded turbulent flows which can become unstable by coupling to a transpiration at the wall. The details of that coupling control the growth rate of the instability, but are relatively unimportant in determining the structure of the eigenfunctions, which are always small perturbations of the neutral inviscid ones.

There is an aspect in which this is not true. The inviscid eigenfunctions depend only of the magnitude $\overline{\alpha}$ of the wavenumber, since the difference between spanwise and oblique waves can be absorbed in Squire's transformation (17). Which wavevector is chosen depends on the transformation properties of the boundary condition (16). We saw for example that in the porous case the boundary condition selects spanwise waves, but other couplings can result in predominantly streamwise, or in essentially isotropic, perturbations. The vertical structure is, on the other hand, only weakly dependent on this.

An intriguing possibility is that the large scales of the logarithmic layer could be explained by such an instability, in which the coupling would be with the active oscillators of the near-wall region. Physically we would be speaking of an interaction between the large outer scales and the near-wall bursting cycle. Although we have seen that the latter can run autonomously, without input from the outer flow, it is probable that it can be modulated by it. The inner layer would, in this view, provide a transpiration boundary condition for the outer one.

Although a lot of work is needed before this conjecture can be tested, there are some hopeful indications. The full eigenvalue problem (13), with homogeneous boundary conditions, can be solved numerically for the mean profile of the $Re_\tau = 590$ channel in [27]. The result is a set of eigenfunctions which depend, assuming $\alpha_3 = 0$, on the streamwise wavelength $2\pi/\alpha_1$. As one would

expect, the eigenfunctions corresponding to short wavelengths are concentrated near the wall and are strongly damped. Longer eigenfunctions penetrate deeper, move faster, and are only weakly damped.

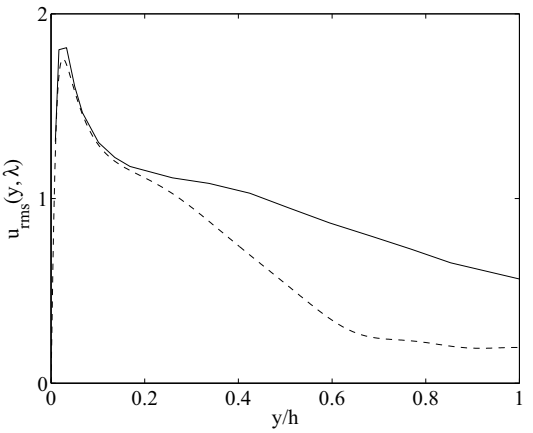

Fig. 8. Fluctuation intensity profiles for the streamwise velocity fluctuations for a wavenumber $k_x/h = 2\pi$. ——— , from the one-dimensional spectra in the channel simulations in [27]; – – – – , from the eigenfunctions of (13). $Re_\tau = 590$. The amplitude scale is arbitrary in both cases.

Since we have energy spectra for the numerical flow in [27], we can compute how the energy contained in a given set of wavelengths depends on the distance to the wall, and compare it to the absolute value of the eigenfunction. They should have similar structures. This is done in figure 8 for $\lambda_x/h = 2\pi$, which is the longest one available in the simulation. They agree fairly well in the logarithmic and wall layers, $y/h < 0.2$, although, given the many approximations involved, this should only be taken as indicative.

5.4 Conclusions

We have reviewed some of the open problems in the physics of the near-wall and logarithmic layers of wall flows. We have seen that the former is an active site for turbulence generation, which can run autonomously and which exports energy to the rest of the flow.

Its structure is reasonably well described in terms of a regeneration cycle involving long streamwise velocity streaks and shorter quasi-streamwise vortices. The cycle seems to be associated to a nonlinear structure, a wavy streak, which has been identified theoretically and observationally, and which we have shown here to be self-sustaining in a severely truncated numerical experiment.

Farther from the wall the structures become larger, with lengths which are much longer than the flow thickness. Even so their spectra scale well with wall

distance, specially for the wall-normal components. We have shown some circumstantial evidence that these very large scales correspond to linear eigenfunctions of the mean velocity profile, weakly damped for homogeneous boundary conditions, but destabilized by their interaction with the active near-wall turbulence cycle.

This work was supported in part by the Spanish CICYT contract PB95-0159, by ONR grant N0014-00-1-0146, and by the Brite EuroMems program, managed by British Aerospace and by Dassault Aviation. Special thanks are due to R.A. Antonia, J. Carlier, R.P. Hoxey, R.D. Moser, A.E. Perry and M. Stanislas for providing their unpublished or hard-to-get data.

References

1. R.A. Antonia, B.R. Pearson: Europhys. Lett. **48**, 163 (1999)
2. R.A. Antonia, M.R. Raupach: Boundary-Layer Meteorol. **65**, 289 (1993)
3. R.F. Blackwelder, H. Eckelmann: J. Fluid Mech. **94**, 577 (1979)
4. K.J. Bullock, R.E. Cooper, F.H. Abernathy: J. Fluid Mech. **88**, 585 (1978)
5. H. Choi, P. Moin, J. Kim: J. Fluid Mech. **262**, 75 (1994)
6. P.G. Drazin, W.H. Reid: *Hydrodynamic Stability* (Cambridge University Press. 1981)
7. T.M. Farabee, M.J. Casarella: Phys. Fluids A **3**, 2410 (1991)
8. H.H. Fernholz, P.J. Finley: Prog. Aerospace Sci. **32**, 245 (1996).
9. M.H. Hites: Scaling of high-Reynolds number turbulent boundary layers in the National Diagnostic Facility. Ph. D. Thesis, Illinois Inst. of Technology (1997)
10. P. Holmes, J.L. Lumley, G. Berkooz: *Turbulence, coherent structures, dynamical systems and symmetry*, (Cambridge Univ. Press. 1996)
11. R.P. Hoxey, P.J. Richards: 'Spectral characteristics of the atmospheric boundary layer near the ground'. In *First UK Wind Engng. Conf. at Cambridge, UK, September 1992*.
12. J.M. Hamilton, J. Kim, F. Waleffe: J. Fluid Mech. **287**, 317 (1995)
13. J. Jeong, F. Hussain, W. Schoppa, J. Kim: J. Fluid Mech. **332**, 185 (1997)
14. J. Jiménez: Phys. Fluids **6**, 944 (1994)
15. J. Jiménez: 'The largest scales of turbulent wall flows'. In *CTR Ann. Res. Briefs*, Stanford Univ. pp. 137–154 (1998)
16. J. Jiménez, P. Moin: J. Fluid Mech. **225**, 213 (1991)
17. J. Jiménez, A. Pinelli: J. Fluid Mech. **389**, 335 (1999)
18. J. Jiménez, M. Uhlmann, A. Pinelli, G. Kawahara: submitted to J. Fluid Mechanics (2000)
19. H.T. Kim, S.J. Kline, W.C. Reynolds: J. Fluid Mech. **50**, 133 (1971)
20. J. Kim, P. Moin, R.D. Moser: J. Fluid Mech. **177**, 133 (1987)
21. K.C. Kim, R.J. Adrian: Phys. Fluids **11**, 417 (1999)
22. M.T. Landahl, E. Mollo–Christensen: *Turbulence and random processes in fluid mechanics* (Cambridge Univ. Press. 1992)
23. C.J. Lawn: J. Fluid Mech. **48**, 477 (1971)
24. M.J. Lee, J. Kim, P. Moin: J. Fluid Mech. **216**, 561 (1990)
25. J.M. Österlund, A.V. Johansson, H.M. Nagib, M.H. Hites: Phys. Fluids **12**, 1 (2000)

26. I. Marusic, J.D. Li, A.E. Perry: 'A study of the Reynolds-shear-stress spectra in zero-pressure-gradient boundary layers'. In *Tenth Australasian Fluid Mech. Conf. at Melbourne, Australia, 11-15 Dec. 1989*. pp. 1.5–1.8.
27. R.D. Moser, J. Kim, N.N. Mansour: Phys. Fluids **11**, 943 (1999)
28. P. Orlandi, J. Jiménez: Phys. Fluids **6**, 634 (1994)
29. A.E. Perry, S. Henbest, M.S. Chong: J. Fluid Mech. **165**, 163 (1986)
30. V.G. Priymak, T. Miyazaki: Phys. Fluids **6**, 3454 (1994)
31. W. Reynolds, W. Tiedermann: J. Fluid Mech. **27**, 253 (1967)
32. S.K. Robinson: Ann. Rev. Fluid Mech. **23**, 601 (1991)
33. C.R. Smith, S.P. Metzler: J. Fluid Mech. **129**, 27 (1983)
34. J.D. Swearingen, R.F. Blackwelder: J. Fluid Mech. **182**, 255 (1987)
35. F. Waleffe: Phys. Fluids **9**, 883 (1997)
36. F. Waleffe: Phys. Rev. Lett. **81**, 4140 (1998)
37. H. Tennekes, J.L. Lumley: *A first course in turbulence* (MIT Press. 1972)
38. S. Toh, T. Itano: preprint, physics/9905012.
39. C.D. Tomkins, Z-C. Lin, R.J. Adrian: Bull. Am. Phys. Soc. **44** (8), 33 (1999)
40. A.A. Townsend: *The structure of turbulent shear flow* (Cambridge U. Press. 1976)
41. M.V. Zagarola, A.J. Smits: Phys. Rev. Lett. **78**, 239 (1997)

Patterning and Transition to Turbulence in Subcritical Systems: The Case of Plane Couette Flow

Paul Manneville[1] and Olivier Dauchot[2]

[1] Laboratoire d'Hydrodynamique, École polytechnique, F-91128 Palaiseau cedex, France
[2] Service de Physique l'État Condensé, Centre d'Études de Saclay, F-91191 Gif sur Yvette cedex, France

Abstract. The problem of the transition to turbulence in subcritical systems is illustrated by experiments and modeling attempts relative to the plane Couette flow. Based on results obtained by the Saclay group (§2), the observed phenomenology points out the role played by coherent structures in the form of streamwise vortices and streaks as nontrivial solutions at finite distance from the basic state. Beyond the formal account in §1, theoretical issues are concretely approached through temporal Lorenz-like and spatio-temporal Swift–Hohenberg-like modeling (§3).

1 Introduction

The transition to turbulence is easier to understand for flows that experience instabilities against infinitesimal perturbations saturating above threshold, and for which the classical tools of weakly nonlinear analysis are appropriate. In particular, the saturation beyond threshold is dealt with by bifurcation theory of linear instability modes using singular perturbation techniques, and space modulations are treated within the multi-scale analysis framework. In these cases, key-words are mainly "envelopes", "patterns", and subsequently "space-time chaos" [1]. The successive bifurcated states involved in the transition then remain in some sense close to each other and "turbulence" sets in at later stages of what can be called a *globally supercritical* scenario. Rayleigh–Bénard convection is a classical example where a physically simple mechanism is at work to produce the instability. However the way is clearly open for more complicated cases involving subcriticality [2].

The most simple example of subcriticality arises in the bifurcation of a single degree of freedom governed by $dA/dt = sA - gA^3$ when $g < 0$ with $s \propto R - R_c$, R being the control parameter and R_c the linear threshold value. The bifurcated solutions then appear below the threshold ($R < R_c$) and are unstable, fig. 1(b), so that the perturbation analysis must be continued in order to get the full branch of nontrivial states. Often these states become stable through a saddle-node bifurcation at $R = R_{sn}$ as in fig. 1(c). Upon varying the control parameter, a competition between different locally stable states at finite distance from each other takes place, instead of the continuous substitution of states that occurs in the supercritical case. A more wildly subcritical case, called *globally subcritical* in

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 58–79, 2001.
© Springer-Verlag Berlin Heidelberg 2001

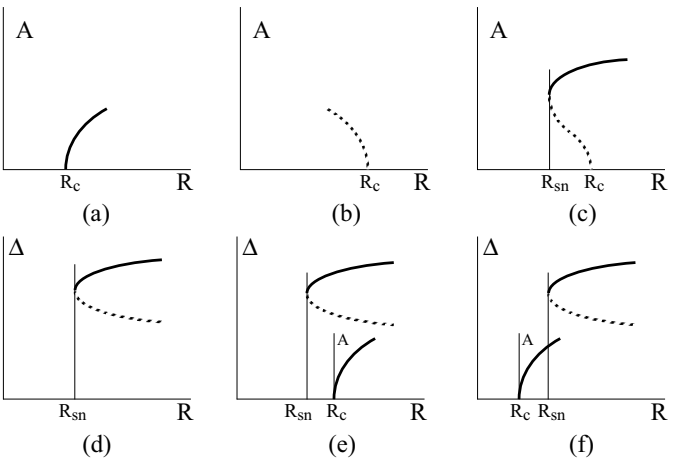

Fig. 1. (a-b) Super/sub-critical bifurcation. (c) Naive saturation in the subcritical case. (d) Globally subcritical bifurcation with a nontrivial branch fully disconnected from an ever locally-stable basic state. (e-f) Competition between a nontrivial branch and a regular one bifurcating from a neutral linear mode (here supercritical, a subcritical bifurcation as in (c) would equally be possible). A represents the amplitude of the bifurcated modes (if any), Δ is the "distance" to the basic state; R_c is the linear instability threshold and R_{sn} the saddle-node threshold at which the subcritical mode (naive or nontrivial) sets in

the following, takes place when the nontrivial solution branch cannot be obtained by perturbation of a neutral linear mode as in fig. 1(d). It can also be that states deriving from neutral linear modes are less relevant to the fully nonlinear problem than those on the disconnected branch as in figs. 1(e,f).

An illustration of the globally subcritical transition to turbulence can be found in open flows. On rather general grounds, it can be shown that the most dangerous linearly unstable modes are waves propagating in the streamwise direction, say x, and independent of the spanwise coordinate z (Squire's theorem), which leads to the classical Rayleigh/Orr–Sommerfeld theory of linear hydrodynamic stability [3]. One has then to distinguish between flows according to whether or not they are mechanically unstable against an inertial Kelvin–Helmholtz-like mechanism resulting from the presence of an inflexion point in the basic velocity profile. The transition to turbulence of inflexional flows, e.g. free shear layers or wakes, takes place at moderate Reynolds numbers and roughly follows a globally supercritical scheme. By contrast, mechanically stable flows are usually unstable against Tollmien–Schlichting (TS) waves arising from subtle viscous effects, but these neutral TS modes typically appear at large Reynolds numbers and may actually not be relevant to the transition scenario. For example, the plane channel Poiseuille flow becomes unstable at $R_c = 5772$ and the TS waves bifurcate subcritically as in fig. 1(c) with $R_{sn} = 2900$, but one can trig-

ger sustained *turbulent spots* down to $R \approx 1000$, which leads us to suspect the existence of a disconnected branch of nontrivial nonlinear states down to such Reynolds numbers [as in fig. 1(e) with a locally supercritical bifurcation replaced by a subcritical one as in fig. 1(c)]. In boundary layers flows (Blasius profile) two kinds of scenarios seem to exist, one globally supercritical involving TS waves, and another globally subcritical involving spots, both merging at later stages when the boundary layer becomes fully turbulent. The case of plane Couette flow (pCf) is the most dramatic since, as in fig. 1(d), it never becomes linearly unstable [4], which makes it particularly attractive for the study of nontrivial nonlinear states, all the more so since perturbations are not carried away by the mean flow, by contrast with Poiseuille and boundary layer flows. The expression "by-pass transition" has been coined to describe scenarios not relying directly on the TS modes and the question has been raised of the mechanisms underlying the growth of perturbations ending in the formation of turbulent spots. In particular, a transient amplification of the perturbations' kinetic energy due to the non-normality of the linear stability operator has been invoked, see e.g. [5], and questioned [6]. Within the "globally subcritical" perspective adopted here, the actual problem is rather the determination of the nature of the nontrivial states involved and, subsequently, the size of the basin of attraction of the basic flow [2].

At odds with the spanwise-independent nonlinear flow patterns deriving from the saturation of TS waves, the coherent structures belonging to the disconnected branch in the scenario sketched above for wall flows are primarily aligned along the streamwise direction. Visually, they present themselves as a spanwise pattern of accelerated and decelerated flow called *streaks* and sustained by a mechanism [7,8] involving the breakdown and reformation of streamwise vortices, but there is no straightforward way to reach them analytically. In the case of pCf, such states have been obtained initially as special nonlinear solutions of the Navier–Stokes (NS) equations by an indirect continuation method starting from a situation where the streamwise vortices are generated by a standard mechanism, Taylor–Couette for Nagata [9] and Rayleigh–Bénard for Busse and Clever [10]. This set of solutions has recently been enriched by Schmiegel [11] who, by direct search, found several other analogous streamwise flow patterns, either steady or time-periodic, of particular interest in view of the discussion of the mechanisms sustaining the streaks [7,8].

In physical space, the phase-space coexistence of different locally stable states characteristic of subcritical systems generically leads to the formation of *domains* homogeneously filled with one of the states in competition and separated from the others by *fronts*. This led Pomeau [12] to analyze the development of spots within laminar flow involved in the transition to turbulence of wall flows as the result of a contamination process and, further to give a specific meaning to the vague concept of *space-time intermittency*. More precisely, the hydrodynamic system was considered to be a distributed two-state (active=turbulent vs. absorbing=laminar) stochastic system with a contamination probability of absorbing states by neighboring active states. In statistical physics, the process

is known as *directed percolation* [13] and serves as a model for the study of epidemics, forest fires, flow through porous media, etc. A transition is observed between transient and sustained activity as the probability is varied through a threshold value, which defines a "universality class" in the sense of critical phenomena, i.e. close to the threshold, statistical properties such as space-time correlations or the activity level display scaling properties that depend on few general characteristics such as the space dimension and the number of absorbing states. Pomeau further conjectured that the transition to turbulence *via* space-time intermittency should belong to the universality class of directed percolation. The perspective thus opened was explored theoretically using models defined in terms of coupled map lattices that helped to understand how spatial coupling could convert transient microscopic chaos into sustained macroscopic space-time chaos with universal features [14]. The concept had interesting applications in several experimental cases where confinement effects reduce the effective space dependence of the physical system considered to one dimension [15] but, somewhat paradoxically, the evidence that it applies to the case it was originally designed for is very limited.

In section 2 below we present experimental results obtained by the Saclay group on the pCf, beginning with illustrations of space-time intermittent regimes, next discussing the dynamics of their ingredients, streaks and streamwise vortices, as observed by triggering or deformation of the basic flow, and concluding by statistical results helping to determine the global stability threshold of the flow. Section 3 is first devoted to the discussion of modeling attempts in the spirit of the Lorenz model of convection appropriate for confined systems. Understanding the behavior of extended systems requires unfreezing the space dependence. This is done by turning to a Swift–Hohenberg-like model, the derivation of which is sketched. A few preliminary simulation results are presented, illustrating the turbulent regime and its decay. Perspectives are discussed in section 4.

2 Experiments on Transitional Plane Couette Flow

The plane Couette flow is the flow established between two parallel plates at distance $2h$ in the $\hat{\mathbf{y}}$ direction, moving along $\hat{\mathbf{x}}$ in opposite directions at speed $\pm U$, see fig. 2(a). The control parameter is the Reynolds number defined as $R = Uh/\nu$, where ν is the kinematic viscosity of the fluid. As already stated, it can be shown theoretically that the laminar velocity profile $\mathbf{v} = u\,\hat{\mathbf{x}}$, with $u = (U/h)\,y$, is stable against infinitesimal perturbations for all R. In spite of the simplicity of the flow, this result has been difficult to obtain and is accordingly rather recent [4]. However, as one could expect, at large R a fully turbulent flow is observed instead of the laminar flow [16]. In fact, it is not easy to produce a high-quality pCf, which explains that most laboratory experiments devoted to the transition regime have been developed during the last decade [17,18].

Here we present experiments performed in Saclay using the set-up sketched in fig. 2(b). It is composed of a plastic loop driven by cylinders. The length of the test section is $L_x \approx 55$ cm, and its width can be varied by changing the loop;

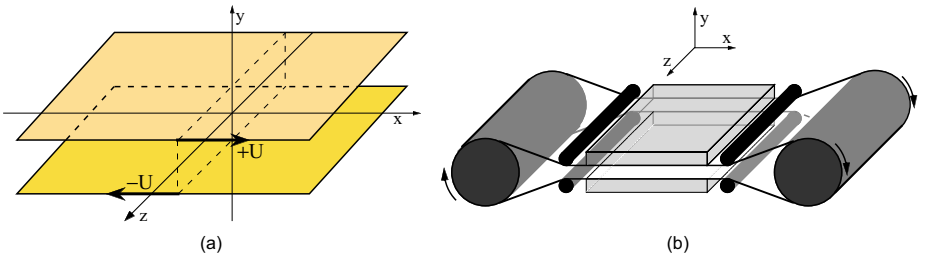

Fig. 2. (a) Geometry of the pCf experiment. (b) Set-up used in Saclay to produce the flow (the actual arrangement is vertical)

most experiments have been performed with $L_z = 25$ cm, some with 12.5 cm. The band is guided by rollers that fix the gap $2h$ from 1.5 mm to 7 mm. The largest aspect-ratios, achieved with the smallest gap, are thus $\Gamma_x \approx 365$ and $\Gamma_z \approx 165$. The working fluid is water at room temperature. The transparency of the plastic belt permits direct optical visualization of eddies (illumination by laser sheets of suspended particles or dye filaments) and quantitative velocity measurements by laser Doppler velocimetry.

2.1 Phenomenology of Transitional Plane Couette Flow

Taking for granted that for sufficiently large R the flow is turbulent, whereas at low R it is surely laminar, one can first examine how the homogeneous turbulent regime breaks down as R is decreased. A typical snap-shot of the flow is displayed in fig. 3. It was taken with the set-up corresponding to the largest available aspect-ratio. A fully turbulent state was prepared and R further reduced down to a value of order 350–380. The so-called "space-time intermittent" regime then obtained presents itself as mixture of patches of turbulent (bright) and laminar (dark) flow. The turbulent fraction, i.e. the relative surface statistically occupied

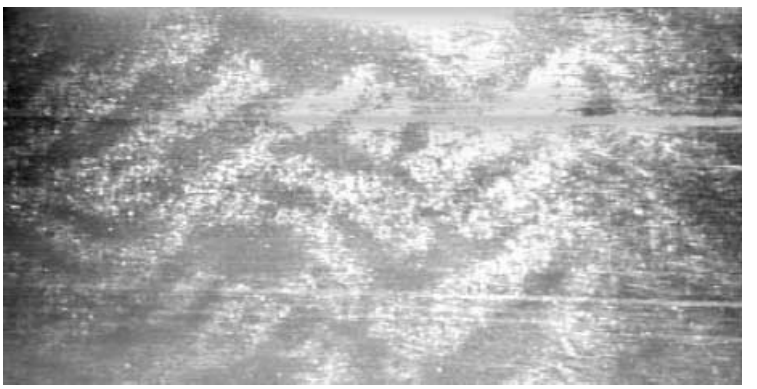

Fig. 3. Turbulent pCf in the space-time intermittent regime at very large aspect ratio

by agitated flow, remains roughly constant at given R and decreases slowly with R (see later). The distribution and shape of turbulent and laminar domains is however not fully random. V-shaped patches or oblique bands of turbulent flow can observed, which is somewhat reminiscent of recent observations of spiral turbulence breakdown in very large aspect-ratio Couette flow between counter-rotating cylinders [19].

At the largest aspect-ratio (achieved with $2h = 1.5$ mm), technical difficulties make it hard to get more quantitative information about the nature of the flow within turbulent patches or the laminar flow between them. Accordingly, from now on we turn to experiments with $2h = 7$ mm. Since the laminar flow is linearly stable, disturbances have to be introduced to nucleate the turbulent regime. In Saclay, repeatable experiments have been performed using a tiny jet shot through the gap as a well-controlled instantaneous and localized perturbation. This perturbation serves as a germ that may abort or develop into a turbulent spot depending on the intensity of the jet and the value of R. A mature turbulent spot observed at $R \simeq 320$ is displayed in fig. 4 where it can clearly be seen that both turbulence inside the spot and the laminar flow around it are not structureless. In particular, more or less regularly spaced streamwise streaky stripes are present all around the spot and especially in the direction of the flow. Laser cuts in the spanwise direction further show [20] that, outside the spot, these structures can be interpreted as paired vortical structure extending over the gap. These structures, though less regular, are still visible in cuts taken inside the spot at its tip, and becomes increasingly disorganized farther inside. The interior of the spot seems filled with similar coherent structures, except that they extend over shorter lengths, and decay and reform on shorter time scales.

Space-time diagrams obtained from spanwise cuts through a spot after nucleation are presented in fig. 5 [21]. The evolution of the streaky stripes is clearly visible. Two cases are represented. After a similar first period of growth, either

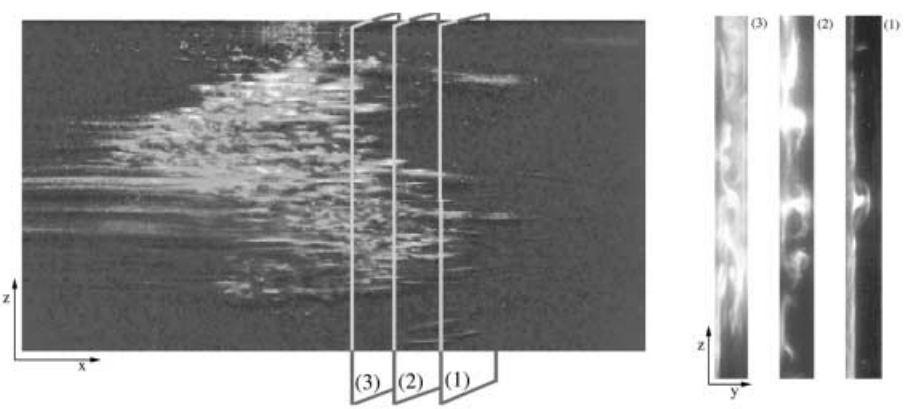

Fig. 4. Structure of the flow around and inside a spot

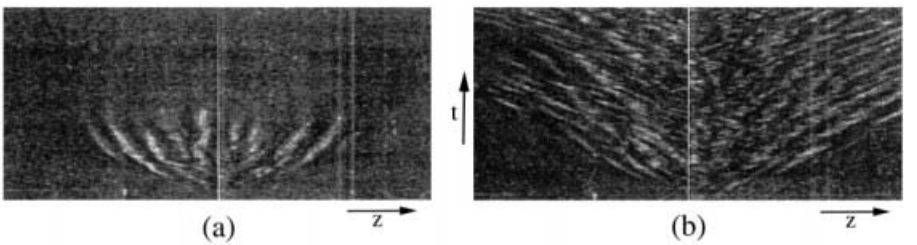

Fig. 5. Space-time diagram illustrating the evolution of streamwise streaky structures: (a) relaxing spot; (b) growing spot; after [21]

the spot stops growing and relaxes, case (a), or it keeps growing, case (b). In the latter case, the turbulent region extends by steady propagation of the stripes, whereas in the former the propagation stops and the structures slowly collapse. Such a collapse is also observed at the late stage of a quench experiment where the Reynolds number is suddenly decreased from a large value corresponding to a fully turbulent regime down to a value just insufficient for sustained space-time intermittent turbulence, see fig. 6.

Since the structures just illustrated seem to be essential ingredients of the transition scenario, it would be advisable to study them in a more or less stabilized form. In the spirit of the continuous-deformation approach to the pCf followed in the theoretical search for non-trivial solutions [9,10,22], experiments have been performed with the basic flow being less and less perturbed by a tiny object of variable size placed in zero-velocity plane midway between the moving walls, either a small bead maintained in the middle of the flow or a thin spanwise wire.

The presence of the bead generates localized pairs of streamwise vortices that further nucleate bursts of space-time activity very similar to that of the jet-triggered spots. The difference is that a sustained temporally intermittent regime can set in since the streamwise vortices reform after the decay of turbulent activity, ready for a new cycle [23], which is not the case for relaxing spots that leave a laminar flow awaiting for extra triggering.

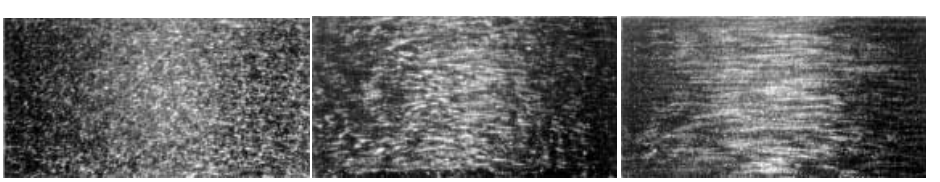

Fig. 6. Three successive snapshots of pCf relaxing toward laminar flow during a quench experiment from $R = 500$ down to $R = 150$; from left to right: before quench time t_0, $t_0 + 5$ sec, and $t_0 + 10$ sec; streamwise and spanwise coordinates are horizontal and vertical, respectively

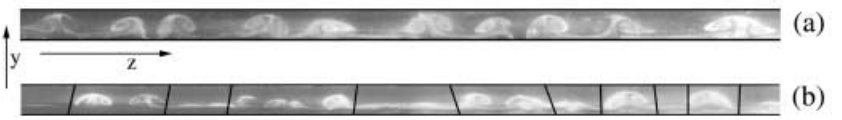

Fig. 7. Bifurcated state (spanwise section) in the region of roughly steady streamwise vortices: (a) for $\rho/h = 0.057$, the streamwise vortices are rather regularly spaced ; (b) for $\rho/h = 0.014$, gaps of laminar flow tend to form between them; after [20]

In practice, it turns out that the study of streamwise vortices is best achieved using the thin wire placed spanwise in the mid-plane. The size of the perturbation so introduced can be measured by the relative radius ρ/h of the wire. Let us consider first the case of a "strong" perturbation with $\rho/h \sim 0.1$. Upon increasing R, a first bifurcation is observed, toward a state made of spanwise-periodic streamwise vortices filling regularly a finite-width band around wire, fig. 7(a). As expected, when ρ/h decreases the threshold of this bifurcation increases (fig. 8), a trend corroborated by the numerical approach of [24] on a slightly different system where the wire is replaced by a ribbon. The latter study further establishes that the bifurcation is subcritical. This is in line with the experimental result that, as the wire is made thinner, the pattern becomes less regular and can be analyzed as resulting from the coexistence of streamwise vortices and gaps where the flow is still in its basic state, fig. 7(b).

The next step is a bifurcation toward a regime where the one-dimensional spanwise pattern of streamwise vortices becomes unsteady in a space-time intermittent fashion. Finally, the streamwise vortices lose their internal coherence and the flow within the unstable band becomes turbulent at $R \simeq 325$ independently of the size of the wire. When decreasing R, the reverse sequence (turbulent band, unsteady and steady streamwise vortices, basic flow) is observed provided that the wire's reduced radius is larger than about 0.02. For the thinnest wires, the turbulent band remains metastable down to $R = 325$ where it decays toward essentially steady streamwise vortices, see fig. 8.

2.2 Statistics and Threshold Determination

The bifurcation diagram in fig. 8 gives the first evidence of a global stability threshold for the *ideal* pCf, i.e. the value R_{g} of the Reynolds number below which the basic flow is the global attractor, the only stable state toward which any initial condition relaxes whatever the nature and duration of transient. In fact, several global thresholds can be defined. The threshold for monotonic stability R_{m}, below which the perturbation's kinetic energy always decays whatever the initial state, is the only one that can be determined analytically. It has been found early to be $R_{\mathrm{m}} = 20.7$ by energy methods [25]. Next, one can define $R'_{\mathrm{g}} = \inf(R_{\mathrm{sn}})$, the minimum of all saddle-node bifurcation points of nontrivial solutions. From the numerical findings obtained so far [9,10,11] one gets $R'_{\mathrm{g}} \sim 125$. In fact, it seems likely that these special solutions, obtained in a minimal-

box context, are attracting only provided that sub-harmonic perturbations (long wavelength modulations) are forbidden, and are unstable saddles otherwise, since stable nontrivial flow has never been observed at such low values of R, at least for the *ideal* pCf to which we restrict ourselves from now on.

Different methods have been used to determine the global threshold experimentally, all converging toward $R_g \simeq 325$. The evolution of the size of triggered spots and that of the turbulent fraction in a quench at given R of homogeneous turbulent flow prepared at larger R are our two main sources of information. Depending on the intensity of the initial perturbation, three possible cases can occur at given R: either the spot relaxes immediately, or it decays after a long transient, or it persists till the end of the experiment (see [26,11] for similar numerical experiments). In the same way, the turbulent fraction in the quench experiment can drop down to zero rapidly, or after a long transient, or stabilize around a specific value that depends on R. In all cases sufficiently many trials have been performed, so as to yield significant statistics. Typical recordings of the relative size of spots (S-traces) and the turbulent fraction in quench experiments (Q-traces) are presented in fig. 9. In both experiments, immediate decay is always observed for $R < R_u \sim 310$ (fig. 9, left) while for $R > R_g$ sustained turbulence is observed in a large fraction of cases (fig. 9, right) but some spots may abort (trace S2). In the intermediate case (fig. 9, middle) long transients are the rule and a turbulent fraction can be measured during the transient but, at the end, the basic state is always recovered.

In practice, the determination of thresholds retains a statistical character and should rather be expressed in terms of statistical ensembles and probabilities, e.g. the probability that long transients exist below R_u is negligible, or the probability that a turbulent spot generated by a given initial perturbation survives at given $R > R_g$ is larger than, say, 50%. In that way, it is legitimate only to define threshold curves for specific perturbations with varying intensities at a given level

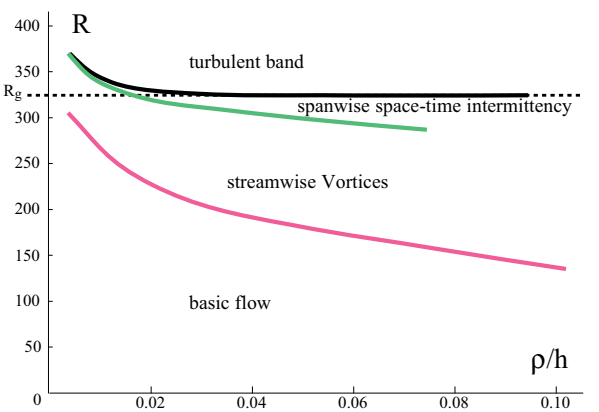

Fig. 8. Bifurcation diagram of the pCf perturbed by the wire; after [20]

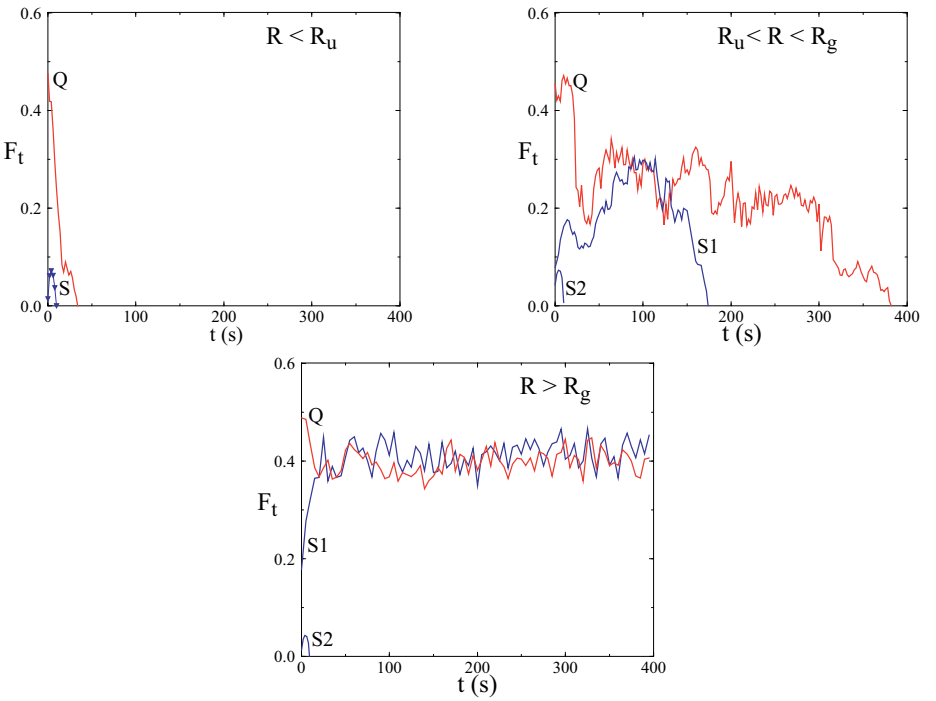

Fig. 9. Turbulent fraction F_t as a function of t in quench (Q-traces) and spot-triggering experiment (S-traces); after [27])

of probability (see [11] for a similar analysis of numerical simulation results). This fact explains most of the discrepancies observed in the literature where different perturbations were used to nucleate spots in different experimental contexts, in the laboratory [17,18] or numerical [28].

For the quench experiment, the measurement of the turbulent fraction yields the bifurcation diagram in fig. 10 where the full line above $R_g \simeq 325$ corresponds to sustained turbulence while the turbulent state along the dashed part is only transient. An analysis of histograms of transient durations τ further shows that for $R < R_g$ the distribution is exponentially decaying with a well-defined average lifetime $\langle\tau\rangle$, i.e. $\mathcal{N}(\tau > \bar{\tau}) \propto \exp(-\bar{\tau}/\langle\tau\rangle)$. Furthermore, as R approaches R_g from below, $\langle\tau\rangle$ is seen to diverge roughly as $1/(R_g - R)$.

The set of experimental results presented above can be interpreted in two ways. The first one relates to the theory of dynamical systems and suggests restricting the discussion to what happens to the state on the nontrivial branch and to understand it as the conversion of a chaotic repellor below R_g into an attractor above by a process akin to a crisis in the vicinity of which chaotic transients with analogous time signatures can be found. The second one, of a more thermodynamic flavor, tries to draw an analogy with first-order (discontinuous) transitions, coexistence, metastability, Maxwell plateaus, and nucleation.

3 Modeling Transitional Plane Couette Flow

3.1 Low-Dimensional Modeling

As far as the theoretical understanding of the transition to turbulence is concerned, the finding of nontrivial nonlinear solutions to the NS equations for the pCf [11] does not solve the problem. We need a dynamical description of the interactions that accounts for the long-term evolution of the system and its bifurcations toward statistically steady but time-dependent regimes. The simplest conceivable modeling is in terms of a differential dynamical system coupling a finite (and small) number of modes in the spirit of temporal chaos studies. Analyzing the process sustaining streaks and streamwise vortices in wall-bounded flows, Waleffe proposed and studied such a model [29] involving the amplitudes of spatially periodic modes with streamwise and spanwise wavevectors (α, γ) fixed at some optimal value derived from the size of the minimal (periodic) box necessary to yield nontrivial behavior in direct numerical simulations [7,8]. His reduced model, obtained by a straightforward Galerkin method plus some simplifications, reads

$$
\begin{aligned}
(d/dt + \kappa_m/R)M &= \sigma_m W^2 - \sigma_m UV + \kappa_m/R, \\
(d/dt + \kappa_u/R)U &= -\sigma_w W^2 + \sigma_u MV, \\
(d/dt + \kappa_v/R)V &= \sigma_v W^2, \\
(d/dt + \kappa_w/R)W &= \sigma_w UW - \sigma_m MW - \sigma_v VW,
\end{aligned}
$$

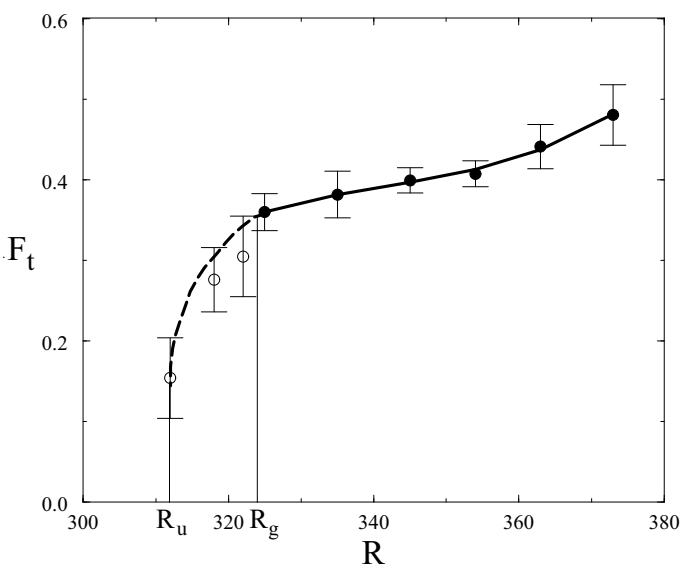

Fig. 10. Turbulent fraction F_t as a function of R in the quench experiment; after [20]

where the κ's and σ's are positive damping and coupling coefficients to be found in [29]. Variable M corresponds to the mean flow, U is the streamwise streak component, V represents the streamwise vortices, and W is a combination of variables accounting for a three-dimensional perturbation. The basic state (BS in fig. 11), corresponding to $M = 1$, all other variables being zero, remains stable for all R.

The bifurcation diagram is slightly more complicated than that in fig. 1(d). For $\alpha = 1.30$, $\gamma = 2.28$, a pair of nontrivial fixed points appears at $R_{\rm sn} = 104.85$. However, by contrast with the one-dimensional case of fig. 1(d) where one of the fixed point is a node (stable) and the other a saddle (unstable), here both fixed points are saddles just beyond the bifurcation. This, which might well be the generic case, is possible since there is sufficient room for other directions of instability, the phase-space being 4-dimensional here. Whereas the lower-branch fixed-point (LB) always remains a saddle, the upper-branch fixed-point (UB) soon becomes an unstable focus at $R = 104.94$ that converts to a stable focus at $R_{\rm H} = 138.06$ through a subcritical Hopf bifurcation giving birth to an unstable limit cycle. As it turns out, despite the fact that the attractors are simply fixed points (BS is a stable node for all R and UB is a stable focus for $R > R_{\rm H}$), the structure of the phase space, further studied in [30], is already quite complicated as seen on the three-dimensional projection of its phase portrait for $R = 160$ in fig. 11.

Depending on the initial conditions, the relaxation toward BS can be immediate (traces a1, a2) or occur only after a long oscillatory transient (traces b1, b2, b3). Trajectories starting in the basin of attraction of UB spiral toward it (traces c1, c2). The shapes of the attraction basins of the competing fixed points, bounded by the stable manifolds of LB and of the unstable limit cycle, are presumably complicated, for example trajectories corresponding to traces b3 and c2 clearly start near the stable manifold of the unstable limit cycle (not shown) but on different sides; b3 returns to BS after having visited the neighborhood of LB, while c2 spirals toward UB. An indirect but more accessible characterization of the attraction basins is by making the statistics of trajectories that relax directly toward one of the stable fixed points or experience long transients depending on the average shape (given direction in phase space) and average amplitude (distance Δ from BS) of appropriately randomized initial conditions, fig. 12.

The limited interest of Waleffe's model lies essentially in the fact that it has only deceptively simple attractors. Other equally simple models have been build, e.g. [31,32], but their relevance to actual pCf is essentially heuristic. By contrast, a systematic attempt has been developed by Eckhardt *et al.* [33,11] using a Galerkin approximation to the Couette flow including a larger set of modes in both the cross-stream and the in-plane directions (36 degrees of freedom before using symmetries to reduce the set to 19, 10, 9 or 4, same general structure as Waleffe's model with linear damping terms and energy-preserving quadratic couplings). A thorough study of the properties of the 9-dimensional model showed the existence of unstable stationary states, unstable periodic or-

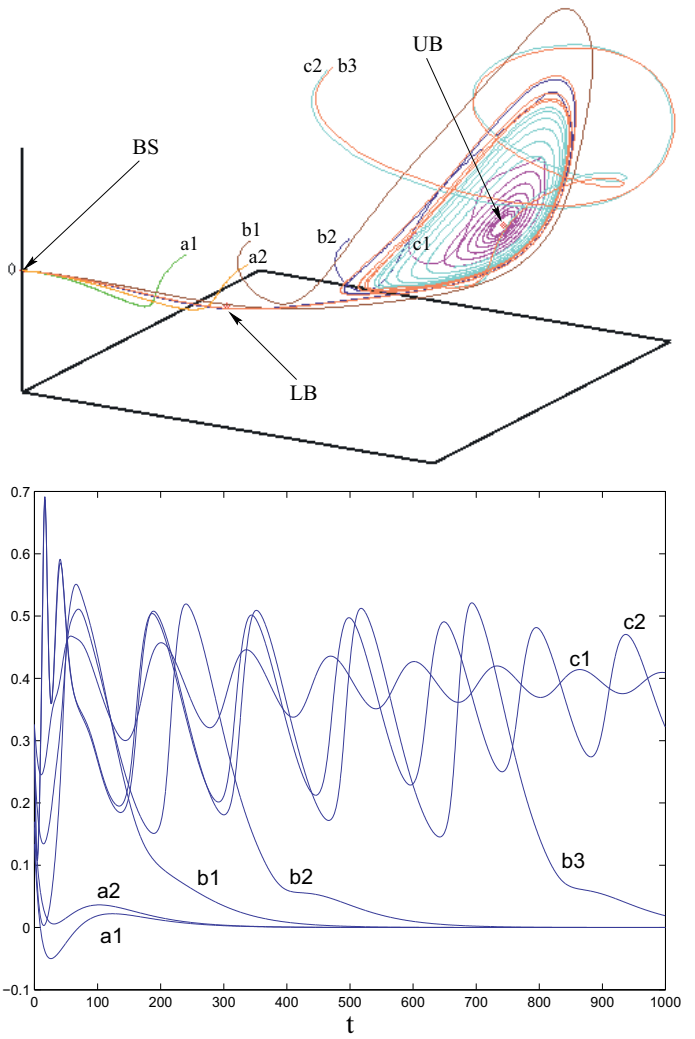

Fig. 11. Above: three-dimensional projection of the phase portrait of Waleffe's model for $R = 160$, BS is the basic state, LB and UB the fixed points on the lower and upper branches, respectively (after [30]); below: time series of one component of the trajectories plotted in the left part

bits relevant to the streaks' breakdown and reformation mechanism, chaotic transients and sustained chaos.

Sharing the same philosophy as the convection model of Lorenz, the attempts briefly reviewed above all work within the framework of dynamical systems and temporal chaos theory, which is valid for confined systems, here at the size of some minimal periodic box. However, the consideration of models which unfreeze

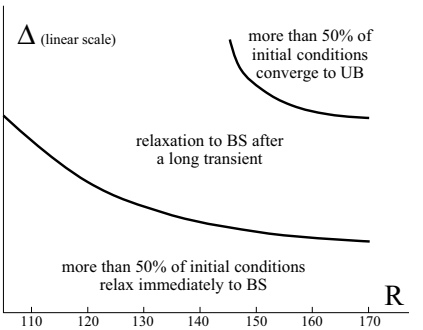

Fig. 12. Statistical definition of the "size" of the attraction basins of the basic state BS (trivial fixed point) and nontrivial attractor UB, after [30].

the space dependence has contributed much to the understanding of weakly turbulent convecting systems. The prototype of such models is that derived by Swift and Hohenberg (SH) who took into account the cross-stream coherence of the flow by severely truncating a Galerkin expansion of the problem but allowed for in-plane modulations in terms of a partial differential equation for the effective local field describing the pattern [34]. The original model had a built-in variational structure appropriate only at the limit of infinite Prandtl number, but already helped in analyzing defects and pattern selection [35]. One step further was taken by including the effects of so-called "drift flows" [36] yielding a reasonable understanding of more exotic regimes [37]. As seen experimentally, transitional pCf takes place at moderate Reynolds numbers and seems to involves a pattern of highly coherent structures in the cross-stream direction, playing the same role as Bénard rolls in convection. In order to go beyond low-dimensional modeling, we present below an attempt where we stress not the fine cross-stream structure as previously done [33,11] but rather the in-plane patterning in line with the approach developed for convection [34,36].

3.2 Modeling "à la Swift–Hohenberg"

To be specific, let us write down the continuity and NS equations

$$\nabla \cdot \mathbf{v} = 0, \tag{1}$$

$$\partial_t \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p + R^{-1} \nabla^2 \mathbf{v} + \mathbf{F}. \tag{2}$$

In a strict parallel with Rayleigh's approach to Bénard convection, the usual realistic no-slip boundary conditions at the moving plates are replaced by stress-free conditions [29]. Since the plate motion is then no longer able to drive the flow, a fictitious streamwise bulk force $\mathbf{F}$ is introduced in equation (2) in the form $\mathbf{F}(y) = F \sin(\beta y)\hat{\mathbf{x}}$ with $\beta = \pi/2$ and $F = \beta^2/R$. The exact solution of (2)

corresponding to the basic state is thus given by $u = u_\mathrm{b} = \sin(\beta y)$.[1] The solution is next sought in a Galerkin expansion of the cross-stream (y) dependence of the hydrodynamic fields using trigonometric functions [38]. The full solution to (2) is then obtained by inserting the velocity and pressure components generated by taking $u = u_\mathrm{b} + u'$ with u' given by

$$u' = U_0 + U_1 \sin(\beta y) + \sum_{k\geq 1} U_{2k} \cos(2k\beta y) + U_{2k+1} \sin((2k+1)\beta y).$$

From the continuity equation (1), it is easily seen that a similar expansion must be taken for the span-wise velocity perturbation component w', whereas the cross-stream component must read

$$v' = V_1 \cos(\beta y) + \sum_{k\geq 1} V_{2k} \sin(2k\beta y + V_{2k+1} \cos((2k+1)\beta y).$$

In the same way, a consistent expression for the pressure can be obtained from the consideration of (2):

$$p' = P_0 + P_1 \sin(\beta y) + \sum_{k\geq 1} P_{2k} \cos(2k\beta y) + P_{2k+1} \sin((2k+1)\beta y).$$

In these expansions it is understood that the fields U_j, V_j, W_j, and P_j depend on x, z, and t. The projection of the continuity and NS equations (1,2) onto trigonometric functions then yields an infinite series of partial differential equations. High order truncations of this system could be considered for better accuracy but the aim is to gain some understanding of mechanisms rather than an accurate representation of a flow that fulfills artificial stress-free boundary conditions, hence modeling by low order truncations. Most of the work to be presented has been performed with a set truncated at first significant order, i.e. involving (U_0, W_0, P_0) and (U_1, V_1, W_1, P_1). Few checks with (U_2, V_2, W_2, P_2) in addition have been performed but will not be reported here.

In Waleffe's picture, the streaks are associated with the velocity components that are uniform in the y-direction, (U_0, W_0), linked together by the projection of the continuity condition onto functions independent of y, which reads:

$$\partial_x U_0 + \partial_z W_0 = 0\,. \tag{3}$$

In the same way the projection of the continuity equation onto $\sin(\beta y)$ yields

$$\partial_x U_1 + \partial_z W_1 = \beta V_1\,. \tag{4}$$

Here only the full expression of the NS equation for U_0 is given, other equations having similar structures. One gets

$$\partial_t U_0 + N_{U_0} = -\partial_x P_0 - \tfrac{1}{2}\partial_x U_1 - \tfrac{1}{2}\beta V_1 + R^{-1}\Delta_2 U_0\,, \tag{5}$$

[1] Notice that in [29], u_b is taken as $\sqrt{2}\sin(\beta y)$, so that R here is larger than that in that reference by $\sqrt{2}$.

where $\Delta_2 = \partial_{xx} + \partial_{zz}$ and

$$N_{U_0} = U_0\partial_x U_0 + W_0\partial_z U_0 + \tfrac{1}{2}U_1\partial_x U_1 + \tfrac{1}{2}\beta V_1 U_1 + \tfrac{1}{2}W_1\partial_z U_1 \qquad (6)$$

accounts for nonlinear advection $\mathbf{v}\cdot\nabla\mathbf{v}$ as expected. The pressure field can further be eliminated by introducing appropriate stream functions and potentials: a stream function Ψ_0 for the flow component independent of y given by $U_0 = -\partial_z\Psi_0$ and $W_0 = \partial_x\Psi_0$, and two fields Ψ_1 and Φ_1 such that $U_1 = \partial_x\Phi_1 - \partial_z\Psi_1$ and $W_1 = \partial_z\Phi_1 + \partial_x\Psi_1$ for the y-dependent part that thus contains both irrotational and rotational contributions. The velocity component V_1 is then derived from equation (4) that now reads:

$$\beta V_1 = \Delta_2\Phi_1\,.$$

The final expression of the model is therefore a system of three partial differential equations for the three two-dimensional fields, Ψ_0, Ψ_1, Φ_1.

The origin of each term in the model is easily identifiable. It can be checked that the advection terms (N_{U_0} and similar) conserve the kinetic energy contained in the perturbation, namely $k = k_0 + k_1 = (U_0^2 + W_0^2) + \frac{1}{2}(U_1^2 + V_1^2 + W_1^2)$. On the right hand side of (5) and other projected NS equations, two kinds of linear terms are to be found, those coming from the linearization of $\mathbf{v}\cdot\nabla\mathbf{v}$ around the basic solution (which give its non-normal character to the linear stability operator) and those expressing the effect of the viscous dissipation.

The model as it stands also exhibits properties of interest with respect to real pCf. First, a classical normal mode analysis, with infinitesimal perturbations varying as $\exp[st + i(\alpha x + \gamma z)]$, shows that the basic flow is linearly stable for all R. The eigen-system indeed reads

$$\begin{aligned}
\left[s + R^{-1}(\alpha^2+\gamma^2)\right]\hat{\Psi}_0 \quad & -i\tfrac{1}{2}\alpha\hat{\Psi}_1 \quad & +i\tfrac{1}{2}\gamma\hat{\Phi}_1 \quad &= 0\,,\\
-i\alpha\hat{\Psi}_0 \quad + \left[s + R^{-1}(\alpha^2+\beta^2+\gamma^2)\right]\hat{\Psi}_1 \quad & & &= 0\,,\\
& & \left[s + R^{-1}(\alpha^2+\beta^2+\gamma^2)\right]\hat{\Phi}_1 \quad &= 0,
\end{aligned}$$

so that the TS mode corresponding to $\hat{\Phi}_1$, with growth rate $s = -(\alpha^2 + \beta^2 + \gamma^2)/R$, is always stable. The other modes (Squire) are mixed and their growth rates are the two roots of

$$s^2 + s(2\alpha^2 + \beta^2 + 2\gamma^2)/R + (\alpha^2+\gamma^2)(\alpha^2+\beta^2+\gamma^2)/R^2 + \tfrac{1}{2}\alpha^2 = 0\,.$$

The product of these roots is always positive and their sum always negative, which yields either two negative roots when $2\alpha^2R^2 < \beta^4$ or two complex roots with negative real part in the opposite case. In fact, the maximum growth rate is $s_{\max} = -\beta^2/2R < 0$ for $\alpha = \gamma = 0$.

Finally, it can be checked that Waleffe's models can be recovered provided that an appropriate ansatz is made for Ψ_0, Ψ_1 and Φ_1. In particular, using his notations and taking

$$\begin{aligned}
\Psi_0 &= -(U/\gamma)\sin(\gamma z) + (A/\alpha)\sin(\alpha x) - B\cos(\alpha x)\cos(\gamma z)\,,\\
\Psi_1 &= (M-1)z - (C/\alpha)\cos(\alpha x) + D\sin(\alpha x)\cos(\gamma z)\,,\\
\Phi_1 &= -(V\beta/\gamma)\cos(\gamma z) - \beta E\cos(\alpha x)\sin(\gamma z)\,,
\end{aligned}$$

where M, U, A, B,... are functions of t only, yields a differential system which, within a first-harmonic approximation, is equivalent to his first model [29] and can be simplified in the same way to give his second model reviewed above.

3.3 Preliminary Results

We present now some results of numerical simulations of the model aiming at a test of its general appropriateness and preparing a super-computer study of larger systems [39]. MatLab running on a personal computer has been used in the simulations. We have considered rectangular domains of size (i) $L_x = 5$, $L_z = 3$, i.e. a confined system in line with the "minimal box" approach, and (ii) $L_x = 32$, $L_z = 16$, a more extended domain where space-time effects are expected. A straightforward pseudo-spectral code appropriate for periodic boundary conditions has been developed. In spite of the fact that the model conserves the energy, some care is required for the choice of the space-time resolution. Computations were found manageable with collocation achieved on a grid with spatial spacing $\delta x = \delta z = 0.125$, which required time-steps down to $\delta t = 0.001$ for numerical stability at the largest R considered.

Simulations corresponding to case (i) have mainly served to make a link with Waleffe's original temporal approach since periodic conditions at a small distance tend to freeze the spatial dependence of the solution. As expected, nontrivial phenomena take place for R of order 200 (140 with Waleffe's scalings). However, no stationary regimes have been found. On the contrary, a sustained chaotic state was obtained for R between 260 and 200 that never settled down to a nontrivial steady or periodic state, whereas long chaotic transients leading directly to the basic state were systematically observed for $R < 195$. The main difference between our model and that of Waleffe lies in the fact that the first-harmonic approximation made to obtain the latter drastically reduces the set of active modes at sufficiently large R. By contrast, in the simulation many more degrees of freedom are kept and their spatial harmonics are no longer neglected or enslaved to the fundamental modes. A direct search for nontrivial solutions of interest and the determination of their basins of attraction as in [11] would also be possible but might not be so rewarding, owing to the prevalence of sustained chaotic states at high R and chaotic transients below.

The most remarkable fact about case (ii) is that nontrivial phenomena take place at much lower Reynolds numbers, which emphasizes the role of effective confinement of periodic boundary conditions placed at small distances, forbidding modulations induced by possible long-wavelength instabilities. In the following, the space average K of the kinetic energy k contained in the perturbation, i.e. $K = \int dx\, dz[(U_0^2 + W_0^2) + \frac{1}{2}(U_1^2 + V_1^2 + W_1^2)]$, as a function of time serves to characterize the evolution of the system; snapshots of the hydrodynamic fields will illustrate its instantaneous state. We report here on quench experiments, fig. 13, the other protocols considered in the experimental section being left for future study.

A statistically stationary turbulent regime was first established for $R = 180$, the state reached at time $t = 200$ serving as initial condition for a set of sim-

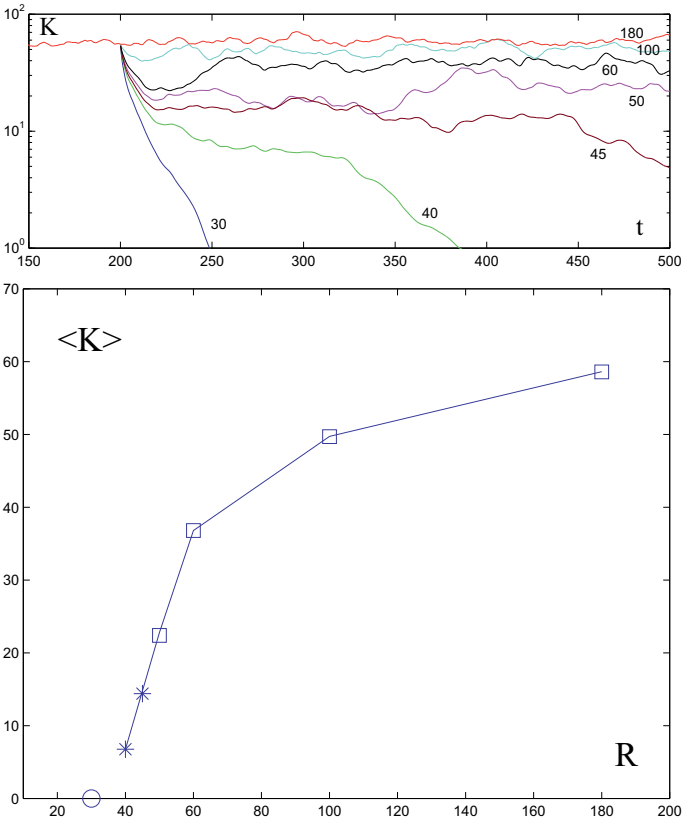

Fig. 13. Above: variation of the total kinetic energy K as a function of time for quench experiments; starting with the state at $t = 200$ of the run at $R = 180$, one decreases R down to the values indicated along the curves; monotonic decay is observed for $R = 30$. Below: time-average of $\langle K \rangle$ as a function of R; the asterisks indicate the values recorded during the plateau part of the transient for $R = 45$ and 40

ulations at smaller values of R. At $R = 100$, a turbulent regime is obtained with slightly smaller time-average of K. Decreasing R down to 50, one seemingly keeps a turbulent regime (only part of the simulation is shown) but with comparatively larger energy fluctuations. For $R = 45$ and 40 the chaotic solution is only transient. It decays past $t \sim 450$ and $t \sim 315$, respectively. The transient is marked by a plateau in the energy variation during which an average can be taken. Finally, for $R = 30$ a direct (without plateau) relaxation of the flow toward the basic state is observed. The "bifurcation diagram" in fig. 13 (right) is analogous to that in fig. 10 corresponding to laboratory experiments in similar conditions.

Typical images of instantaneous states are displayed in fig. 14 for $R = 50$ in the sustained turbulent regime ($t = 750$, left column) and for $R = 40$ ei-

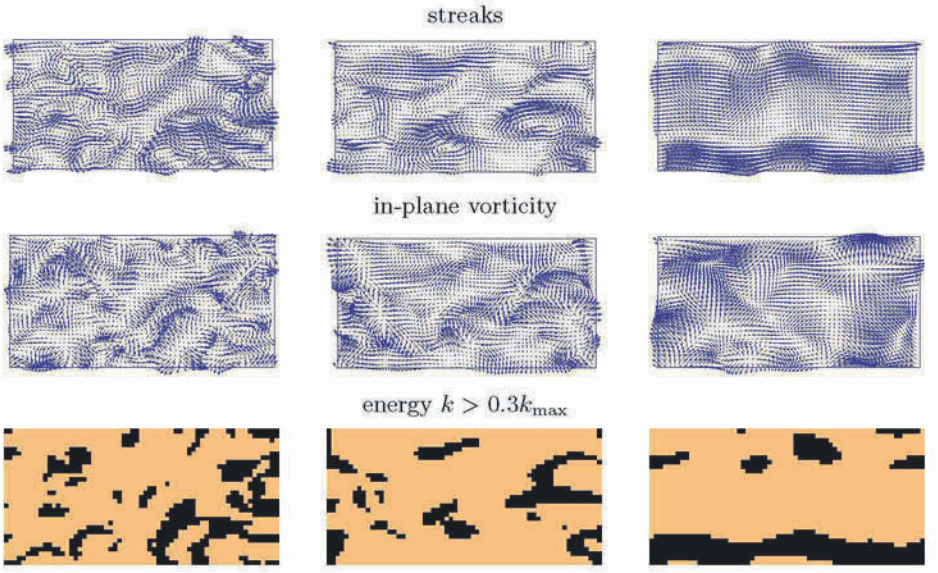

Fig. 14. Snapshots of the simulated flow for $R = 50$, $t = 750$ (left), and for $R = 40$, $t = 250$ (middle) and $t = 400$ (right); see text for details

ther during the chaotic transient ($t = 250$, middle column) or during the terminal relaxation stage ($t = 400$, right column). The streak component of the perturbation (U_0, W_0) is given in the top row. The middle row displays the in-plane vorticity vector field with components $\zeta_x = (\beta W_1 - \partial_z V_1)\cos(\beta y)$ and $\zeta_z = (\partial_x V_1 - \beta U_1)\cos(\beta y)$ further evaluated in the plane $y = 0$ and from which one should infer the pattern of vortices (streamwise wherever $|\zeta_z| \ll |\zeta_x|$). The total kinetic energy contained in the perturbation is depicted in the last row after thresholding at 0.3 of its maximum value. In the fully turbulent state, at $R = 180$ or 100, the energy distribution turns out to be much more intermittent than for $R = 50$, and the resulting in-plane small-scale structure observed calls into question the reliability of the lowest order model with limited y-resolution. (This restriction applies only to the case of extended systems since confinement effects contribute to kill the finest structures, so that reliable conclusions can be drawn at higher R within the framework of a low-dimensional approach.) As seen in fig. 14, this is no longer the case when $R \leq 50$ for which streaks or streamwise vortices can be easily identified, either directly or from the localization of their energy content, and especially in the decaying stage of transient chaos. The average anisotropy of the vortices seems however less marked than in the experiment, which might be related to the fact that the stress-free boundary conditions used in the model are less stringent than the realistic no-slip conditions.

4 Perspectives

As a typical example of globally subcritical system we have considered the case of pCf which, when ideal, is linearly stable for all Reynolds numbers. Experimentally, patterns of streamwise vortices have been shown to play a role in the transition as nontrivial solutions at finite distance from the basic state, the distinctive feature of subcriticality. In fact, pCf is not a pathological system. The most similar one is of course the cylindrical Couette flow in counter-rotating configuration, especially with respect to the decay of spiral turbulence [40,19]. But, in view of the ubiquity of streamwise vortices, streaks, and their cycle of breakdown and regeneration [7,8], the study of transitional plane Poiseuille flow or Blasius boundary layer flow should also benefit from the enlightenment gained in considering the simplest possible case of wall-bounded flow. In this respect, semi-realistic modeling temporal attempts [29,33] already bring interesting information about the nature of transients [30], the change from transient to sustained chaos [33,11], and the statistical significance of thresholds.

Complementary to this approach in line with "minimal box" ideas, we have stressed on the need for a space-time unfreezing of the pattern and an examination of its consequences. A SH-like model has thus been derived [38]. Preliminary simulation results [39] are encouraging in that they display properties rather similar to those in the experiments. As far as extended systems are of interest, the limitations of our lowest order modeling are now clearly identified. Concerning the numerics, simulations of systems sufficiently large to permit the coexistence of laminar and turbulent domains typical of spot dynamics in the space-time intermittent regime would be welcome. So, within its range of reliability, our model as it stands seems to deserve both large-scale super-computer simulations and further analytical study. Improved modeling along the same lines could relax some pathological features arising from the stress-free boundary conditions used (the same underestimation of the relevant range of Reynolds number as for Rayleigh's treatment of convection, and the same spurious presence of a neutral vorticity mode accounted for by Ψ_0).

At any rate, the recourse to models seems compulsory, especially in order to understand the nature of the bifurcated state and the status of transients in very large transitional systems, and to test the appropriateness of the assumptions implicit in the space-time intermittent scenario [12].

In contrast with convection-like globally supercritical systems, wall flows generically experience a subcritical transition to turbulence that leaves many open questions. In this respect, we hope to have shown the paradigmatic position of pCf and to have proposed a useful tool to attack this delicate problem.

Acknowledgements

The authors want to thank S. Bottin and F. Daviaud for their contribution to the experimental work, L. Tuckerman, Y. Pomeau, and H. Chaté for interesting discussions about the theoretical interpretation.

References

1. M.C. Cross, P.C.Hohenberg, "Pattern formation outside equilibrium", Rev. Mod. Phys. **65** (1993) 851–1112.
2. O. Dauchot, P. Manneville, J. Phys. II France **7** (1997) 371–389.
3. P.G. Drazin, W.H. Reid, *Hydrodynamic stability, Cambridge University Press* (Cambridge University Press, 1981)
4. V.A. Romanov, Funkt. Anal. Ego Priloz. **7** (1970) 62–73.
5. L.N. Trefethen, A.E. Trefethen, S.C. Reddy, T.A. Driscoll, Science **261** (1993) 578-584.
6. F. Waleffe, Phys. Fluids **7** (1995) 3060-3066.
7. J.M. Hamilton, J. Kim, F. Waleffe, J. Fluid Mech. **287** (1995) 317–348.
8. J. Jiménez, A. Pinelli, J. Fluid Mech. **389** (1999) 335–359.
9. M. Nagata, J. Fluid Mech. **217** (1990) 519–527.
10. R.M. Clever, F.H. Busse, J. Fluid Mech. **234** (1992) 511–527, and "Bifurcation sequences in problems of thermal convection and of plane Couette flow", in: J. Grue *et al*, eds., *Waves and Nonlinear Processes in Hydrodynamics* (Kluwer, Dordrecht, 1996).
11. A. Schmiegel, "Transition to turbulence in linearly stable shear flows", Marburg, Dissertation, 1999 (`http://archiv.ub.uni-marburg.de/diss/z2000/0062/`).
12. Y. Pomeau, Physica D **23** (1986) 3–11. See also P. Bergé, Y. Pomeau, Ch. Vidal, *L'espace chaotique* (Hermann, Paris, 1998) Ch. IV.
13. W. Kinzel, "Directed Percolation", dans G. Deutscher *et al.*, eds., *Percolation Structures and and Processes*, Annals of the Israel Phys Soc., vol.5 (1983).
14. H. Chaté, P. Manneville, "Spatio-temporal intermittency", in *Turbulence, a tentative dictionary*, P. Tabeling & O. Cardoso, eds. (Plenum Press, 1995).
15. F. Daviaud, "Experiments in 1D turbulence", in *Turbulence, a tentative dictionary*, P. Tabeling & O. Cardoso, eds. (Plenum Press, 1995).
16. H. Reichardt, "Gesetzmäßigkeiten der geradlinigen turbulenten Couetteströmung", Mitteilungen AVA Göttingen **22** (1959).
17. N. Tillmark, P.H. Alfredsson, F. Fluid Mech. **235** (1992) 89–102.
18. F. Daviaud, J. Hegseth, P. Bergé, Phys. Rev. Lett. **69** (1992) 2511–2514.
19. A. Prigent, O. Dauchot, "Barber-pole turbulence in large aspect-ratio Taylor–Couette flow", Saclay preprint, Phys. Rev. Lett. submitted (2000); "Laminar-turbulent coexistence following a subcritical transition to turbulence", at 8th European Turbulence Conference, ETC8, Barcelona, June 2000.
20. S. Bottin, O. Dauchot, F. Daviaud, P. Manneville, Phys. Fluids **10** (1998) 2597–2607.
21. O. Dauchot, F. Daviaud, Phys. Fluids **7** (1995) 335–343.
22. A. Cherhabili, U. Ehrenstein, J. Fluid Mech. 342 (1997) 159-177.
23. S. Bottin, O. Dauchot, F. Daviaud, Phys. Rev. Lett. **79** (1997) 4377–4380.
24. D. Barkley, L.S. Tuckerman, Phys. Fluids **11** (1999) 1187–1195.
25. D.D. Joseph, *Stability of fluid motions. I*, Springer Tracts in Natural Philosophy, vol. 27 (Springer, Berlin, 1976), see p. 178ff.
26. A. Schmiegel, B. Eckhardt, Phys. Rev. Lett. **79** (1997) 5250–5253.
27. S. Bottin, F. Daviaud, P. Manneville, O. Dauchot, Europhys. Lett. **43** (1998) 171–176.
28. A. Lundbladh, A.V. Johansson, J. Fluid Mech. **229** (1991) 499–516.
29. Waleffe F., Phys. Fluids **9** (1997) 883–900.
30. O. Dauchot, N. Vioujard, Eur. Phys. J. B **14** (2000) 377–381.

31. Th. Gebhardt, S. Grossmann, Phys. Rev. E **50** (1994) 3705–3711.
32. J.S. Baggett, L.N. Trefethen, Phys. Fluids **9** (1997) 1043–1053.
33. B. Eckhardt, A. Mersmann, Phys. Rev. E **60** (1999) 509–517.
34. J. Swift, P.C. Hohenberg, Phys. Rev. A **15** (1977) 319.
35. H.S. Greenside, W.M. Coughran Jr., N.L. Schryer, Phys. Rev. Lett. **49** (1982) 726–729.
36. P. Manneville, J. Physique **44** (1983) 759–765.
37. M. Bestehorn, M. Fantz, R. Friedrich, H. Haken, C. Pérez-García, Z. Phys. B –Condensed Matter **88** (1992) 93–94.
38. P. Manneville, F. Locher, C.R. Acad. Sci. Paris **328**, série IIb (2000) 159–164.
39. P. Manneville, "Modélisation de la transition vers la turbulence de l'écoulement de Couette plan", in *Rencontre du Non-linéaire 2000* (Paris Onze Éditions, Univ. Paris-Sud, Orsay, 2000); also "Modeling transitional plane Couette flow", at 8th European Turbulence Conference, ETC8, Barcelona, 2000.
40. C.D. Andereck, S.S. Liu, H.L. Swinney, J. Fluid Mech. **164** (1986) 155.

Effects of the Upstream Conditions on the Mean Statistics of Turbulent Boundary Layers with Zero Pressure Gradient

Luciano Castillo[1], Gunnar Johansson[2], and Erik Eisenmen[1]

[1] Rensselaer Polytechnic Institute
Department of Mechanical and Aeronautical Engineering
Troy, NY 12180 USA
[2] Chalmers University of Technology
Department of Thermo and Fluid Dynamics
SE-412 96 Göteborg, Sweden

Abstract. A low Reynolds number experiment in a zero pressure gradient, ZPG, turbulent boundary layer has been carried out using the LDA technique and similarity analysis of the RANS equations. This experiment seeks to investigate the effect of the upstream conditions on the downstream development of the mean flow and turbulent quantities. It was found that, by fixing the upstream conditions, the mean velocity profiles, and Reynolds stresses tend to collapse even tough the Reynolds number, R_θ, varies from about 700 up to about 5,500. However, whenever the upstream conditions are changed for a given experiment the data will not collapse, thus it is Reynolds number dependent. But this dependence will be shown to come from the variation of the upstream conditions and not from the local Reynolds number. Three different wind tunnel speeds were investigated: 5 m/s, 10 m/s and 20 m/s respectively for a given trip wire size and location.

1 Introduction

The initial effort by Prandtl [11], who successfully developed the boundary layer theory, was indeed a significant contribution to the turbulent field. Later, von Karman [14], continued the efforts in boundary layers and suggested that boundary layers should be independent of the Reynolds number, and that a universal logarithmic function describing the overlap region for both inner and outer mean velocity profiles should exist. This pioneer work was based on the assumption that there is a single velocity scale (the friction velocity, u_*) describing turbulent boundary layers and that a constant stress layer exists.

Thus, it follows that the mean velocity profiles, and Reynolds stresses should scale with the friction velocity, u_*, and must be independent of the streamwise direction, x. Therefore, any quantity must be independent of the Reynolds number and should collapse with just the friction velocity, u_*. Clauser [6] extended the work of Millikan [10] and von Karman [14] to include boundary layers with pressure gradient, PG. He defined an equilibrium boundary layer as one where

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 80–91, 2001.
© Springer-Verlag Berlin Heidelberg 2001

the outer velocity deficit given by,

$$\frac{U_\infty - U}{u_*} = f_o(\overline{y}) \tag{1}$$

should collapse as a single curve, and the pressure parameter given by

$$\beta = \frac{\delta_*}{\rho u_*^2}\frac{dP_\infty}{dx} = constant \tag{2}$$

must be a constant for a given flow. So both conditions must be satisfied for an equilibrium flow to exist. However, these conditions as well as any implications from a single velocity scale assumption of the classical theory were not satisfied [1], [3]. Thus, he concluded that equilibrium flows are very special flows which are very difficult to achieve in experiments.

Townsend [13] used similarity principles for the RANS equations but over constraint the problem by assuming that a single velocity scale for boundary layers must exist. Castillo [1] showed in great details the lack of collapse on the mean profiles and that boundary layer flows grow differently depending on their initial conditions.

More recently, Castillo [1], and Castillo and George [3] showed, using similarity analysis, that the proper outer velocity scale in turbulent boundary layers is the free stream velocity, U_∞, and for the Reynolds shear stresses, $U_\infty^2 d\delta/dx$. These results apply for both ZPG and PG boundary layers. However, they were not able to completely collapse the data with just, U_∞, and attributed the lack of collapse to the local Reynolds number. The implications of their analysis was that boundary layers are characterized by two velocity scales, thus it depends on the Reynolds number. But only in the limit as the Reynolds number approach to infinity the boundary layer becomes completely independent of the Reynolds number. It then follows that all mean statistics profiles should collapse, but only in the limit. This is in fact the Asymptotic Invariance Principle, AIP, proposed by George [7].

Recently, Castillo [2], and Castillo et al. [5] used similarity principles to derive the Zagarola and Smits [16] scaling, $U_\infty(\delta_*/\delta)$, and were able to show that only three velocity profiles exist in boundary layers: one for adverse pressure gradient, one favorable pressure gradient, and one for zero pressure gradient. This conclusion is consistent with the similarity analysis results given in terms of the pressure parameter, Λ, of Castillo and George [3].

Moreover, Castillo [2], and Castillo et al. [5] showed that the Reynolds number dependence observed in the flow was primarily due to the changes of the upstream conditions, and not to the local Reynolds number dependence as suggested by George and Castillo [8] for the ZPG boundary layers. For example, figure 1 shows the ZPG outer deficit profiles of Smith and Walker [12] (left side) and the data from Wieghardt [15] (right side) for ZPG boundary layer. Notice the Reynolds number dependence of the Smith and Walker data. However, the profiles are approaching toward an asymptotic profile as the Reynolds number increases. This observation is consistent with the AIP, and the results

from George and Castillo [8]. On the other hand, the velocity deficit profiles of Wieghardt collapse as a single curve. Notice that the Smith and Walker data has a range of Reynolds number between 3,000 up to about 48,000 and the data from Wieghardt from about 1,000 up to about 17,000. However, the greatest dependence on the Reynolds number occurs from about 1,000 up to about 6,000. So any variation with Reynolds number on the flow should have been seen more in the profiles of Wieghardt than in the Smith and Walker profiles. The question is: why does the data from Smith and Walker vary with Reynolds number while the data from Wieghardt collapses as a single curve?

Looking carefully at the Smith and Walker data the wind tunnel speed changes for each profile from about 44 m/s up to 92 m/s, whereas the data from Wieghardt was measured at a fixed wind tunnel speed (fixed upstream conditions) of 33 m/s and the downstream distance, x, was varied. Furthermore, Smith and Walker also varied the downstream distance. Therefore, it is clear from these two plots that the variation with Reynolds number of the Smith and Walker data is due primarily to the changes in the upstream conditions. Castillo et al. [5] showed a large number of examples of the effects of the upstream conditions on the velocity deficit profiles including boundary layers with pressure gradient.

Although they showed the influence of the upstream conditions on the velocity deficit profiles, they did not show if by fixing the upstream conditions, the turbulence statistics will collapse, or how the turbulence quantities will be affected by the conditions imposed upstream of the flow. Thus, the purpose of this investigation is to study how the upstream conditions, such as the wind tunnel speed and size of the trip wire could affect the velocity deficit profiles and turbulence quantities in a ZPG turbulent boundary layer.

2 Description of the Experiment

A low Reynolds number experiment based on the momentum thickness, R_θ, from about 700 up to about 5,000 was designed and carried out to test how the upstream conditions may influence other turbulent quantities such as the velocity fluctuations and Reynolds stresses. The ZPG turbulent boundary layer considered was 2-D, incompressible, steady state on the mean, and the flat plate was smooth. The details of the experiment are described below.

2.1 The Test Section

The measurements were carried out in the wind-tunnel L2 at the department of Thermo and Fluid Dynamics at Chalmers University of Technology. This wind tunnel has a measuring section 3 m long, 1.8 m wide and 1.25 m high. The corners are provided with fillets, slightly decreasing in size in the downstream direction to compensate for boundary layer growth on the wind tunnel walls. The wind-tunnel is of conventional closed-loop design, equipped with turning

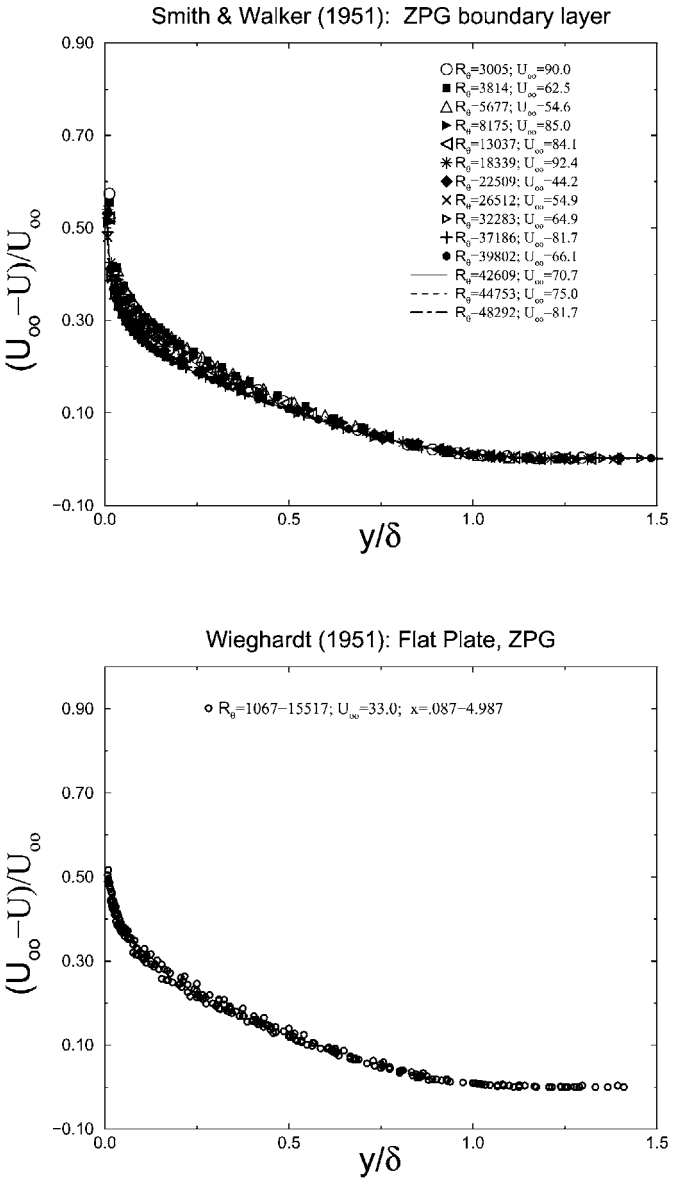

Fig. 1. Velocity deficit profiles in ZPG boundary layer normalized by U_∞ and δ_{99}.Above is the data of Smith and Walker, below is the data of Wieghardt.

vanes in all four corners with a number of honeycombs and screens, and the contraction ratio is 5.6:1. The free stream turbulence level is 0.1%.

A flat plate 2.5 m long, 1.25 m wide and 5 mm thick, was specially manufactured. It was mounted horizontally in the wind-tunnel with the tip of its

nose 200 mm downstream of the start position of the measuring section of the wind tunnel. It was located at a distance of 540 mm from the top wall of the wind-tunnel, i. e., above mid-height, because we had not access to an optical probe with a focal length more than 600 mm. The boundary layer was tripped using a wire with a diameter of 2 mm (d_o= 2 mm). It was positioned across the plate at a distance of 150 mm (x_o= 150 mm) from the leading edge of the plate. The flat plate had a specially designed nose to ensure that separation will not take place. This was also confirmed both computationally and experimentally, the details are given in reference [4].

2.2 The Measuring Probes

For the two dimensional flow, two LDA probes were used. The nearly horizontal probe was used to emit blue and green light from an Argon-ion laser. Two expanders with expansion ratio of 1.94:1 were used with one focusing beam expander with an expansion ratio of 1.55:1 and a focal length of 1200 mm. The top probe was used to collect the light scattered from the measuring control volume created by the horizontal probe. With this side-scattering arrangement the measuring control volume was nearly spherical with a diameter of about 58μm. Optical access to the measuring section was created by fitting two glass windows into the wind tunnel on the side and top walls.

The two probes were connected by aluminum beams to a traverse table allowing simultaneous motion of the two probes. The traverse distance in the main stream direction was from 0.8 m up to 2 m measured from the leading edge of the plate and about 80 mm in the vertical direction from the surface of the plate.

The scattered light collected by the vertical probe was processed by two Dantec BSA processors using the Burstware 3.22 program. With the present geometry of the beams the normal velocity component is measured in a direction slightly side-ways, i. e., it is also somewhat sensitive to the transverse velocity component. The beam angles were measured within 1% thus, the mean velocity in the normal direction could be corrected with very good accuracy (based on the assumption that the transverse mean velocity is zero). Such a correction is, however, not possible for the turbulence quantities if not also the third velocity component is measured.

A potentially worse problem is created if the measurement of the normal velocity component is also sensitive to the velocity in the main stream direction. This problem was avoided by performing test measurements very close to the wall ($y^+ \approx 2$), where the vertical velocity component is very close to zero to satisfy the boundary conditions at the wall. The probe was turned around its own axis until no mean velocity in the vertical direction could be detected.

For measurements very close to the wall, in the viscous sub-layer, accurate determination of the distance between the measuring control volume and the wall is crucial. This was done by traversing the measuring control volume down to the wall and seeking the vertical position where the scattering from the wall itself was maximum. The correct vertical position can be determined to within about

10 μm in this way, corresponding to about 0.3 viscous wall units ($y^+ \approx 0.3$) at a free stream speed of 10 m/s. This process was repeated for each axial position.

3 The Similarity Analysis

A brief description of the similarity analysis will be given here. However, detailed information of the method can be found in references [1], [8], [3].

- **Similarity Analysis:**
  The outer scales of the turbulent boundary layer equations must be determined from an equilibrium similarity analysis of the governing equations, and not chosen *a priori* or using dimensional analysis. George and Castillo [8] applied this concept to the RANS equations in order to determine the mean velocity and Reynolds shear stress scales in the outer turbulent boundary layer as U_∞ and $U_\infty^2 d\delta/dx$ respectively. By matching the Reynolds shear stress in inner and outer variables in the limit as $\delta^+ \to \infty$ they found that $u_*^2 \sim U_\infty^2 d\delta/dx$.

- **Similarity Solution Forms:**
  The basic assumption is that it is possible to express any dependent variable, in this case the outer deficit velocity, $U - U_\infty$, and the outer Reynolds shear stress, $< uv >$, as a product of two functions, thus,

$$U - U_\infty = U_{so}(x) f_o(\overline{y}, \delta^+; *) \tag{3a}$$

$$- < uv > = R_{so}(x) r_o(\overline{y}, \delta^+; *) \tag{3b}$$

  where U_{so} and R_{so} are the outer velocity scale and the outer Reynolds shear stress scales respectively and depend on x only. Note that the outer velocity scale, U_{so}, and the outer Reynolds shear stress scale, R_{so}, are unknown at the moment and *must* both be determined from the boundary layer equations. The arguments inside the similarity functions f_o and r_o represent the outer similarity length scale, $\overline{y} = y/\delta_{99}$, the Reynolds number dependence, $\delta^+ = \delta u_*/\nu$, and any possible dependence on the upstream conditions, *, respectively.

- **Asymptotic Invariance Principle: AIP**
  This principle means that in the limit as the $Re \to \infty$ the boundary layer equations become independent of the Reynolds number, therefore, any function or scale must also be independent of δ^+ as well, but only in the limit as $Re \to \infty$. Since in the infinite *local* Reynolds number limit, the outer equations become independent of δ^+, so must the solutions to them. Hence, in this limit equations 3a and 3b must also become independent of Reynolds number (as required by the AIP); i.e.,

$$f_o(\overline{y}, \delta^+; *) \to f_{o\infty}(\overline{y}) \tag{4a}$$

$$r_o(\overline{y}, \delta^+; *) \to r_{o\infty}(\overline{y}) \tag{4b}$$

  as $\delta^+ \to \infty$. The subscript $o\infty$ is used to distinguish these infinite Reynolds number solutions from the finite Reynolds number profiles used before (equations 3a and 3b). Castillo et al [5] showed that $* = f(\delta_*/\delta)$ and when

ZPG and PG boundary layers are normalized by the Zagarola/Smits scaling, $U_\infty \delta_*/\delta$, all the Reynolds number dependence and effects of the upstream conditions, *, vanish from the outer deficit profiles. Thus, asymptotically the profiles are independent of the upstream conditions as well. Furthermore, note that $\delta_*/\delta = f(\delta^+)$, thus in accordance with the AIP, $\delta_*/\delta \rightarrow constant$ which implies that $* \rightarrow constant$ consistent with equations 4a and 4b.

4 Results

The results from the experiment will be presented in this section. The goal is to show how the mean velocity profiles and the Reynolds stresses are affected by the upstream conditions. The three cases investigated were at wind tunnel speeds of 5 m/s, 10 m/s and 20 m/s, however due to space limitation only the results at 5 and 10 m/s will be presented here. The diameter of the trip wire was 2 mm (d_o= 2 mm) and it was located at a distance of 150 mm (x_o =15 mm) from the leading edge of the plate. In fact, a Reynolds number based on these parameters could easily be calculated to describe the conditions at the leading edge. For example, the Reynolds number based on the wire location, $R_{xo} = U_\infty x_o/\nu$, or based on the wire diameter, $R_{do} = U_\infty d_o/\nu$ could easily be calculated. Recall that the wind tunnel speed and the trip wire size represent the upstream conditions for this experiment, other possible conditions could exist such as suction or blowing.

The velocity deficit profiles are normalized by the free stream velocity, U_∞, and the boundary layer thickness, δ_{99}. The Reynolds stresses profiles were normalized by U_∞^2 and δ_{99}. The results for the two wind tunnel conditions (5 m/s and 10 m/s) are given in the following order: top left plot are the velocity deficit profiles, and below are the crosstream component of the Reynolds stresses. At the top of the right hand side are the streamwise component of the Reynolds stresses profiles, and just below are the Reynolds shear stresses.

Figures 2,3 show all the profiles described above for the wind tunnel speed of 5 m/s. Note that all profiles collapse reasonable well. However, notice that the normal stresses profiles closest to the trip wire (x_o = 0.8, 0.9 m) tend to be more sensitive to the trip wire. Therefore, they don't collapse as well as the profiles far from the trip wire. But, this observation is clearly seen on the streamwise component of the turbulent field. The range of Reynolds number based on momentum thickness, R_θ, for this wind tunnel speed is from about 700 up to about 1,600 and the streamwise positions vary from 0.8 m up to 2.0 m downstream, having a total of 12 profiles. In general we can say that the turbulence quantities and the mean deficit profiles collapse reasonable well for the 5 m/s wind tunnel speed.

Figures 4,5 show the profiles at a wind tunnel speed of 10 m/s in the same order as in the previous figures. Notice that all of the profiles collapse very well even the streamwise component of the the normal stress. The range of Reynolds number based on the momentum thickness, R_θ, is from about 1,400 up to about 3,000. It is in this range where the strongest dependence of Reynolds number

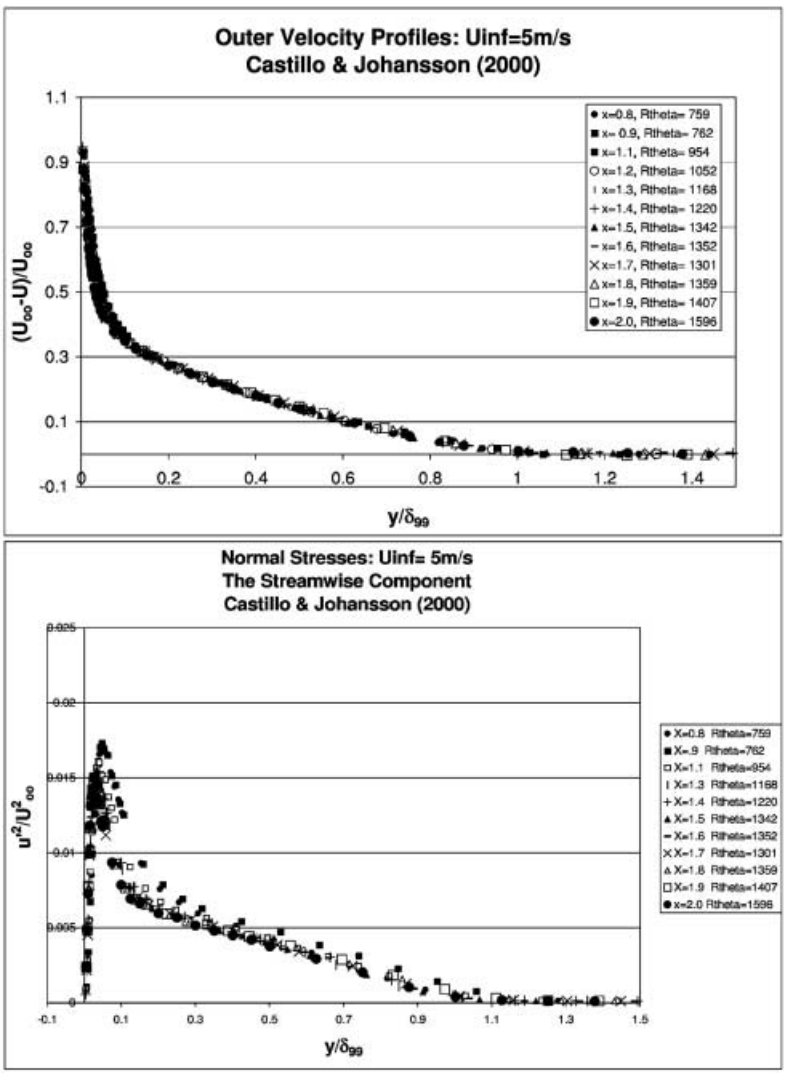

Fig. 2. Profiles in outer variables at a wind tunnel speed of 5 m/s. The Reynolds number vary from $759 \leq R_\theta \leq 1,596$ and the streamwise distance from $0.8 \leq x \leq 2$ m.

has been observed previously. However, these profiles clearly show a very nice collapse but for fixed upstream conditions. Comparing the collapse of the profiles at 5 m/s and 10 m/s it can be observed that as the wind tunnel speed increases the profiles become less sensitive to the trip wire, thus the data collapse better.

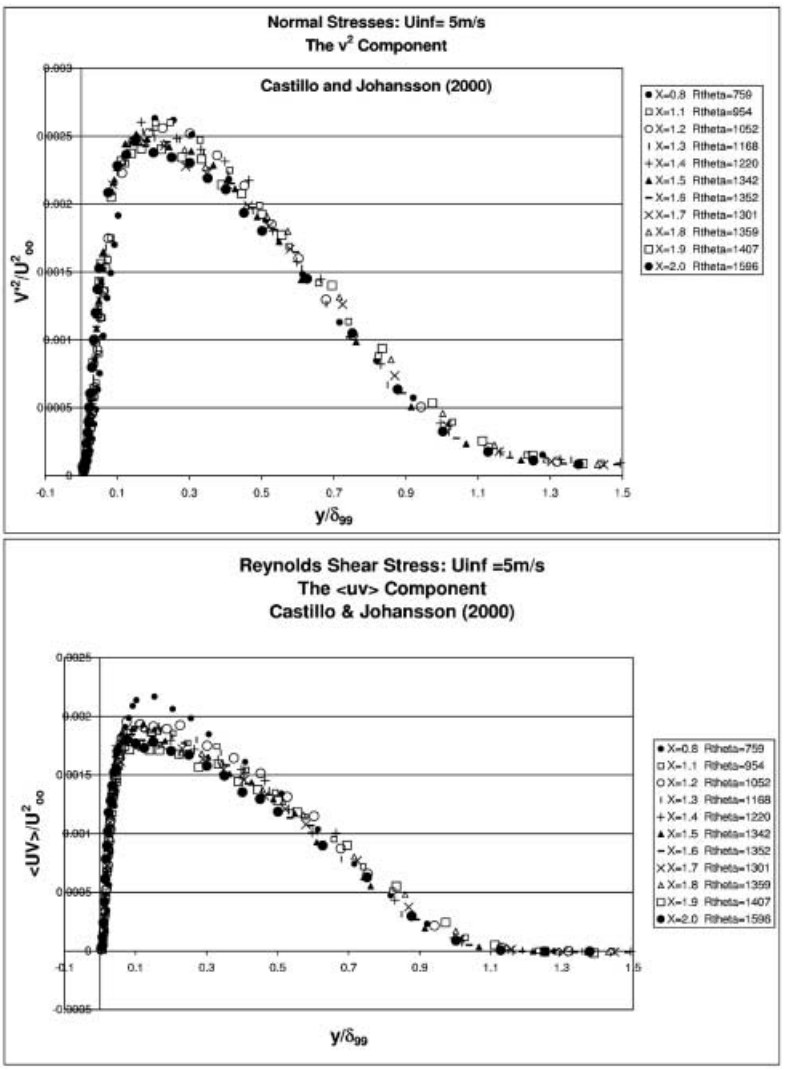

Fig. 3. Profiles in outer variables at a wind tunnel speed of 5 m/s. The Reynolds number vary from $759 \leq R_\theta \leq 1,596$ and the streamwise distance from $0.8 \leq x \leq 2$ m.

5 Conclusion

It is clear that the Reynolds number dependence observed in some ZPG boundary layers is primarily due to the changes in the upstream conditions, consistent with the observations of Castillo et al [5]. It was shown that by fixing the upstream conditions the velocity deficit profiles and Reynolds stresses profiles tend to collapse. For a wind tunnel speed of 5 m/s the first few profiles close to the trip wire (x = 0.8, 0.9 m) of the streamwise Reynolds stresses deviate slightly

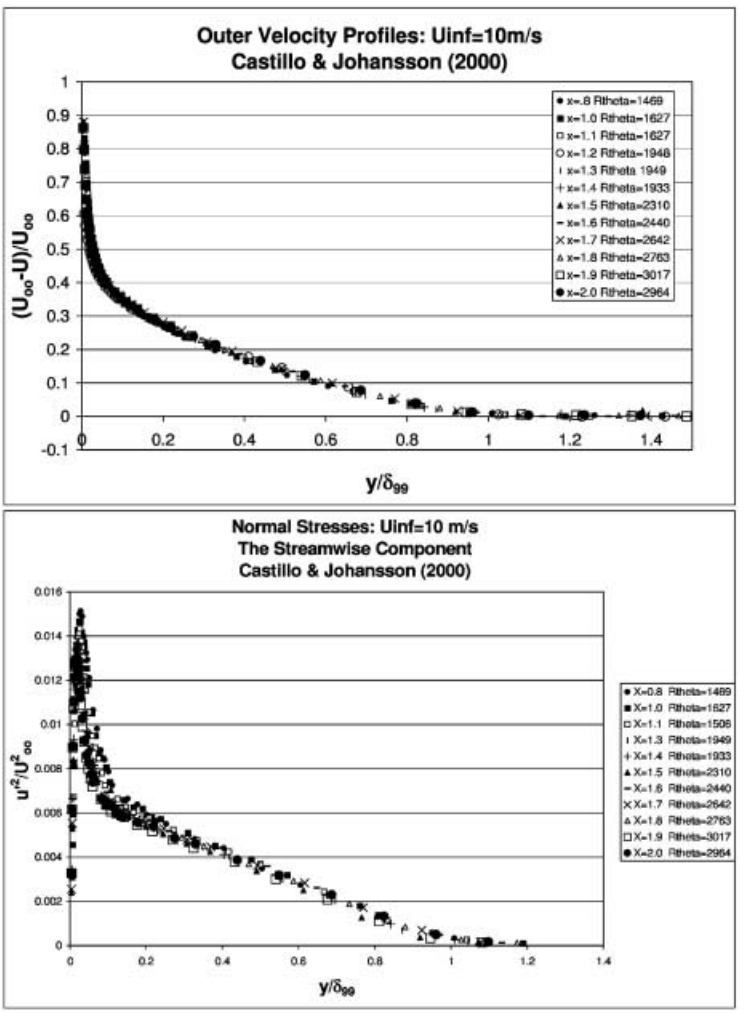

Fig. 4. Profiles in outer variables at a wind tunnel speed of 10 m/s. The Reynolds number vary from $1469 \leq R_\theta \leq 3,017$ and the streamwise distance from $0.8 \leq x \leq$ 2m.

from the collapsed profiles. However, this deviation was very weak at 10 m/s. The profiles for the normal Reynolds stresses collapsed all the way at 10 m/s free stream velocity and fairly well at 5 m/s. The Reynolds shear stress collapsed fairly well, the deviations being more pronounced for the lower free stream velocity. However, if accurate values of $d\delta/dx$ are obtained then the collapse of the data will improve drastically. As shown in the similarity analysis the Reynolds shear stress component, $< uv >$, scale with $U_\infty^2 d\delta/dx$ and not just with U_∞^2 as the u^2, v^2 and w^2 components.

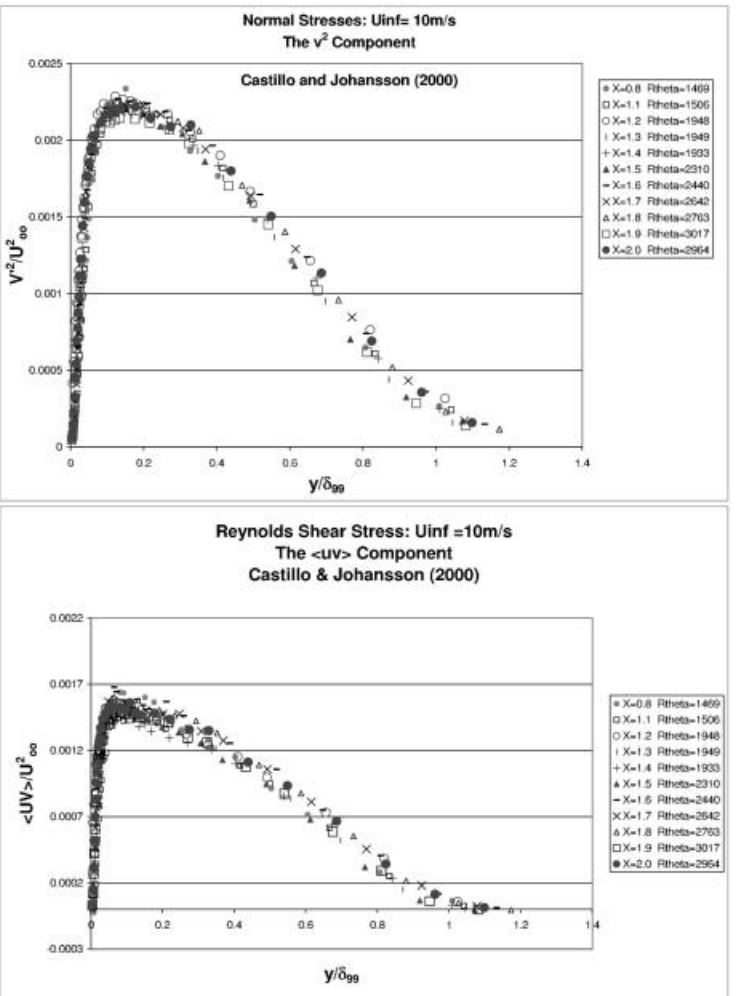

Fig. 5. Profiles in outer variables at a wind tunnel speed of 10 m/s. The Reynolds number vary from $1469 \leq R_\theta \leq 3,017$ and the streamwise distance from $0.8 \leq x \leq 2$m.

Clearly, the data indicates that the effects of the upstream conditions on the downstream flow decreases as the wind tunnel speed increases. Furthermore, the fact that the upstream conditions play an important role in the flow development in the downstream region means that flow control due to upstream manipulations may be a feasible and powerful tool. This can be very useful for boundary layers near separation.

Acknowledgement

The authors are very thankful to Professor William K. George for his valuable comments. Also, thanks go to David J. Walker for making the corrections in legends and converting the excel figures into eps format.

References

1. L. Castillo: Similarity Analysis of Turbulent Boundary Layers. Phd Thesis, State University of New York at Buffalo, New York (1997).
2. L. Castillo: 'Application of Zagarola/Smits Scaling in Turbulent Boundary Layers with Pressure Gradient'. *Advances in Fluids Mechanics 3, Montreal, Canada, May 24–26 2000*, ed. by M. Rahman and C.A. Brebbia pp. 275–288.
3. L. Castillo and W.K. George: 'Boundary Layers with Pressure Gradient: Similarity of the Velocity Deficit Region'. 38^{th} *AIAA Ann. Mtg.* paper $\#2000-0913$ Reno, NV, Jan 10-13, (2000).
4. L. Castillo and G. Johansson: An Experimental Study of the Upstream Conditions on a Canonical Turbulent Boundary Layer, to be summitted to Physiscs of Fluids, August 2000.
5. L. Castillo, D. Walker and M. Wosnik: 'The Effect of the Upstream Conditions on the Mean Velocity Deficit of Turbulent Boundary Layers'. *Fluids 2000 Conference and Exhibit* paper $\#AIAA2000-2309$. Denver, Colorado, June 19-22, (2000).
6. F.H. Clauser: The Turbulent Boundary Layer, *Adv. Appl. Mech.*, IV, 1-51 (1954).
7. W.K. George: 'Some New Ideas for Similarity of Turbulent Shear Flows'. *Turbulence, Heat and Mass Transfer*, ed. by K. Hanjalic and J.C.F. Pereira, (Begell House, NY 1995)
8. W.K. George and L. Castillo: Zero-Pressure Gradient Turbulent Boundary layer, *Applied Mechanics Reviews*, vol 50, no.12, part 1, December (1997).
9. W.K. George and L. Castillo: 'Boundary Layers with Pressure Gradient: Another Look at the Equilibrium Boundary Layer'. *Near Wall Turbulent Flows*, ed. by R.M.C. So. et al. (Elsevier, New York 1993) pp 901–910
10. C.M. Millikan: A Critical Discussion of Turbulent Flows in Channels and Circular Tubes *Proc. 5th Int. Congr. Appl. Mech.*, 386-392, Wiley, NY (1938).
11. L. Prandtl: *Ü*ber Flüssigkeitsbewegung bei sehr kleiner Reibung, Verhandlung des III Intern. *Math. -Kongresses, Heidelberg, 1904, pp.484-491, Leipzig.* See also *Vier Abhandlungen zur Hydrodynamik u. Aerodynamik, Göttingen, (1927)*, Edwards (1943).
12. D.W. Smith and J.H. Walker: Skin-friction measurements in incompressible flow, *NACA Rept.* R 26 (1959).
13. A.A Townsend: The Structure of Turbulent Shear Flows, *Cambridge University Press*,Cambridge (1956).
14. T. von Karman: Mechanishe *Ä*ehnlichkeit und Turbulenz *Nachr. Ges. Wiss.*, Göttingen, pp 68 (1930).
15. K. Wieghardt: Über die Wandschubspannung in turbulenten Reibungsschichten bei veränderlichem Aussendruck, *Kaiser Wilhelm Institut für Strömungsforschung*, No.U&M-6603, Göttingen, Germany (1943).
16. M.V. Zagarola and A.J. Smits: Mean-Flow Scaling of Turbulent Pipe Flow, *J. Fluid Mech.* vol. 373, pp.33-79 (1998).

Evolution of the Large-Scale Structures in the 'Far-Field' of Turbulent Shear Flows

Dan Ewing

McMaster University, Hamilton ON, L8S 4L7

Abstract. One of the most widely accepted ideas in turbulence theory is that the large-scale structures in turbulent shear flows forget how they are generated so that all flows of a particular type asymptotically evolve to a single *universal* state. It is shown here, using the temporally evolving wake as an example, that the governing equations for the two-point velocity correlations have self-similar solutions indicating the large-scale structures asymptotically evolve in an equilibrium manner. It is further shown these equations do not have unique solutions so the notion of a single 'universal' solution is not necessarily consistent with the governing equations.

1 Introduction

It is well known that the single-point moments, such as the mean velocity or the turbulent Reynolds stress, have self-similar profiles in the far field of many turbulent free-shear flows. It is also often conjectured that the turbulent structures in these flows eventually forget how they were generated and decay to a generic structure as the flow evolves. Thus, all flows of a particular type, such as plane wakes or jets, asymptotically evolve to an 'universal' solution characterized by a single length and velocity scale[1]. This idea has gained wide acceptance and is often consider a 'working rule' of turbulence[2].

Wygnanski et al.[3] performed one of the most extensive tests of this conjecture in the early 1980's by measuring the wake behind a series of different plane wake generators that had the same drag. They found that the velocity and turbulence stress profiles in the individual wakes were self-similar. However, the Reynolds stress profiles from the different wakes did not collapse to a single 'universal' profile as expected. George[4,5] later showed that the governing equations for the self-similar mean velocity and Reynolds stress profiles in this (and other) flows included a growth rate dependent parameter. Thus, the equations do not have a single unique solution and wakes with different growth rates, such as those measured by Wygnanski et al., *should* have different self-similar profiles. Both George and Wygnanaski et al. proposed that the differences in the observed growth rates and self-similar profiles occurred because the large-scale structures in the flow do not forget how they were generated and carry information about the initially conditions of the flow into its far field. Neither, demonstrated that this conjecture was consistent with equation that govern the development of the large-scale structures in turbulent flows.

Ewing [6] later demonstrated that the governing equations for the two-point velocity correlations, which are a measure of the large-scale turbulent structures,

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 92–102, 2001.
© Springer-Verlag Berlin Heidelberg 2001

have self-similar solutions in several different free-shear flows. He also found that, like the equations for the single-point moments, the equations for the self-similar solutions of the two-point correlations included a growth rate dependent parameter and thus *did not have an unique solution.* This paper gives an overview of these results using the temporally evolving wake as an example and discusses their physical implications. It is also shown that the similarity analysis can be used to demonstrate that the evolution of the flow in the temporally evolving plane wake, often used in Direct Numerical Simulations of plane wake flows, is asymptotically analogous to the evolution of the flow in the self-propelled wake.

2 Temporally Evolving Plane Wake

The geometry of the temporally evolving plane wake, shown in figure 1, is similar to the self-propelled plane wake left behind a two-dimensional body moving through stationary air. In the temporally evolving wake, though, the evolution of the wake in the mean-flow direction is ignored and it is assumed that the wake is homogeneous in this direction. This simplification makes it much easier to perform Direct Simulations of this flow[7] than the self-propelled wake. As it will be shown later the first order equations of these flows are asymptotically the same (to the level of the two-point correlations) so using the temporally evolving wake to study the spatially evolving or self-propelled is analogous to using grid turbulence in a wind tunnel to study decaying isotropic turbulence.

Following the approach outlined by George[5], Moser et al [8] have shown previously that the governing equations for the single-point moments have self-similar solutions. For example, the averaged moment equations given by

$$\frac{\partial \delta U_1}{\partial t} = -\frac{\partial \overline{u_1 u_2}}{\partial x_2} \tag{1a}$$

$$\rho \int_{-\infty}^{\infty} -\delta U_1 dx_2 = \dot{m}, \tag{1b}$$

have self-similar solutions of the form

$$\delta U_1 = U_s(t) f(\eta) \tag{2a}$$

$$\overline{u_1 u_2} = R^{12}(t) r_{12}(\eta), \tag{2b}$$

where $\eta = x_2/\delta(t)$, if

$$U_s(t) \propto 1/\delta(t) \tag{3a}$$

$$R^{12} \propto U_s \frac{d\delta}{dt} \ . \tag{3b}$$

Here, $-\delta U_1 = U_\infty - U_1$, u_1 and u_2 are the fluctuation components in the streamwise, x_1, and the cross-stream, x_2, directions (cf. figure 1) and $\dot{m}$ is the mass deficit of the wake.

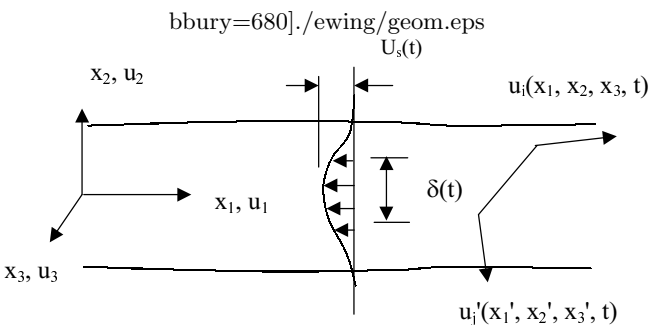

Fig. 1. Geometry of the temporally evolving wake.

In the temporally evolving wake, the governing equations for the two-point single-time velocity correlations [9] in the temporally evolving wake can be reduced to

$$\frac{\partial \overline{u_i u_j'}}{\partial t} + (\delta U_1 - \delta U_1') \frac{\partial \overline{u_i u_j'}}{\partial r_1} =$$

$$-\frac{1}{\rho}\left[\frac{\partial}{\partial r_1}\left(\overline{pu_j'}\delta_{i1} - \overline{p'u_i}\delta_{j1}\right) + \frac{\partial \overline{pu_j'}}{\partial x_2}\delta_{i2} + \frac{\partial \overline{p'u_i}}{\partial x_2'}\delta_{j2} + \frac{\partial}{\partial r_3}\left(\overline{pu_j'}\delta_{i3} - \overline{p'u_i}\delta_{j3}\right)\right]$$

$$-\frac{\partial}{\partial r_1}\overline{(u_1 - u_1')u_i u_j'} - \frac{\partial \overline{u_2 u_i u_j'}}{\partial x_2} - \frac{\partial \overline{u_2' u_i u_j'}}{\partial x_2'} - \frac{\partial}{\partial r_3}\overline{(u_3 - u_3')u_i u_j'}$$

$$-\overline{u_j' u_2}\frac{\partial \delta U_1}{\partial x_2}\delta_{i1} - \overline{u_i u_2'}\frac{\partial \delta U_1'}{\partial x_2'}\delta_{j1} + \nu\left(2\frac{\partial^2}{\partial r_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_2'^2} + 2\frac{\partial^2}{\partial r_3^2}\right)\overline{u_i u_j'}, \quad (4)$$

where the unprimed variables are evaluated at one point in the wake, (x_1, x_2, x_3, t), and the primed variables are evaluated at a second arbitrary point at the same time, (x_1', x_2', x_3', t) as shown figure 1. Here, $r_1 = x_1 - x_1'$ and $r_3 = x_3 - x_3'$ are the separation distance in the statistically homogeneous x_1 and x_3 directions, respectively.

It is hypothesized these equations have self-similar solutions of the form

$$\overline{u_i u_j'} = Q^{i,j}(t) q_{i,j}(\xi, \eta, \eta', \zeta, *) \quad (5a)$$

$$\overline{pu_j'}/\rho = \Pi^{,j}_{(1)}(t)\pi^{(1)}_{,j}(\xi, \eta, \eta', \zeta, *) \quad (5b)$$

$$\overline{p'u_i}/\rho = \Pi^{i,}_{(2)}(t)\pi^{(2)}_{i,}(\xi, \eta, \eta', \zeta, *) \quad (5c)$$

$$\overline{(u_\beta - u_\beta')u_i u_j'} = T^{i,j}_{(\beta)}(t) tt^{(\beta)}_{i,j}(\xi, \eta, \eta', \zeta, *) \quad (5d)$$

$$\overline{u_i u_2 u_j'} = T^{i,j}_{(2)}(t) tt^{(2)}_{i,j}(\xi, \eta, \eta', \zeta, *) \quad (5e)$$

$$\overline{u_i u_2' u_j'} = T^{i,j}_{(2p)}(t) tt^{(2p)}_{i,j}(\xi, \eta, \eta', \zeta, *)\ , \quad (5f)$$

where $\xi = r_1/\delta_1(t)$, $\eta = x_2/\delta(t)$, $\eta' = x_2'/\delta(t)$, and $\zeta = r_3/\delta_3(t)$ are the similarity variables. Substituting these solutions into the governing equations yields

$$\left[\frac{dQ^{i,j}}{dt}\right] q_{i,j} - \left[\frac{Q^{i,j}}{\delta_1}\frac{d\delta_1}{dt}\right] \xi \frac{\partial q_{i,j}}{\partial \xi} - \left[\frac{Q^{i,j}}{\delta}\frac{d\delta}{dt}\right] \left(\eta \frac{\partial}{\partial \eta} + \eta' \frac{\partial}{\partial \eta'}\right) q_{i,j} - \left[\frac{Q^{i,j}}{\delta_3}\frac{d\delta_3}{dt}\right] \zeta \frac{\partial q_{i,j}}{\partial \zeta}$$

$$+ \left[\frac{U_s(t) Q^{i,j}}{\delta_1}\right] \{f(\eta) - f(\eta')\} \frac{\partial q_{i,j}}{\partial \xi} = -\left(\left[\frac{\Pi^{,j}_{(1)}}{\delta_1}\right] \frac{\partial \pi^{(1)}_{,j}}{\partial \xi} \delta_{i1} - \left[\frac{\Pi^{i,}_{(2)}}{\delta_1}\right] \frac{\partial \pi^{(2)}_{i,}}{\partial \xi} \delta_{j1} \right.$$

$$\left. + \left[\frac{\Pi^{,j}_{(1)}}{\delta}\right] \frac{\partial \pi^{(1)}_{,j}}{\partial \eta} \delta_{i2} + \left[\frac{\Pi^{i,}_{(2)}}{\delta}\right] \frac{\partial \pi^{(2)}_{i,}}{\partial \eta'} \delta_{j2} + \left[\frac{\Pi^{,j}_{(1)}}{\delta_3}\right] \frac{\partial \pi^{(1)}_{,j}}{\partial \zeta} \delta_{i3} - \left[\frac{\Pi^{i,}_{(2)}}{\delta_3}\right] \frac{\partial \pi^{(2)}_{i,}}{\partial \zeta} \delta_{j3} \right)$$

$$- \left[\frac{T^{i,j}_{(1)}}{\delta_1}\right] \frac{\partial tt^{(1)}_{i,j}}{\partial \xi} - \left[\frac{T^{i,j}_{(2)}}{\delta}\right] \frac{\partial tt^{(2)}_{i,j}}{\partial \eta} - \left[\frac{T^{i,j}_{(2p)}}{\delta}\right] \frac{\partial tt^{(2p)}_{i,j}}{\partial \eta'} - \left[\frac{T^{i,j}_{(3)}}{\delta_3}\right] \frac{\partial tt^{(3)}_{i,j}}{\partial \zeta}$$

$$- \left[\frac{Q^{2,j} U_s}{\delta}\right] q_{2,j} \frac{df}{d\eta} \delta_{i1} - \left[\frac{Q^{i,2} U_s}{\delta}\right] q_{i,2} \frac{df}{d\eta'} \delta_{j1}$$

$$+ \nu \left(2\left[\frac{Q^{i,j}}{\delta_1^2}\right] \frac{\partial^2}{\partial \xi^2} + \left[\frac{Q^{i,j}}{\delta^2}\right] \frac{\partial^2}{\partial \eta^2} + \left[\frac{Q^{i,j}}{\delta^2}\right] \frac{\partial^2}{\partial \eta'^2} + 2\left[\frac{Q^{i,j}}{\delta_3^2}\right] \frac{\partial^2}{\partial \zeta^2} \right) q_{i,j} \tag{6}$$

These equations have self-similar solutions if all the time-dependent terms in square brackets are proportional. Physically, this implies that all the energy transfer processes in the flow are in equilibrium as the flow evolves in time.

It is straightforward to show this condition is satisfied if

$$\delta_1 \propto \delta \propto \delta_3 \tag{7}$$

so that the characteristic size of turbulent motions grow at the same rate in all directions as the wake grows. Similarly, the constraints for the two-point correlations are given by

$$Q^{1,1} \propto Q^{3,3} \propto Q^{1,3} \propto Q^{3,1} \propto Q^{2,3} \propto Q^{3,2} \propto U_s^2 \tag{8a}$$

$$Q^{1,2} \propto Q^{2,1} \propto Q^{2,2} \propto U_s(t) \Delta\delta/\Delta t = \beta U_s^2(t) \tag{8b}$$

$$\Pi^{,1}_{(1)} \propto \Pi^{1,}_{(2)} \propto \Pi^{,3}_{(1)} \propto \Pi^{3,}_{(2)} \propto U_s^2 \Delta\delta/\Delta t = \beta U_s^3 \tag{8c}$$

$$\Pi^{,2}_{(1)} \propto \Pi^{2,}_{(2)} \propto \beta^2 U_s^3 \tag{8d}$$

$$T^{i,j}_{(1)} \propto T^{i,j}_{(2)} \propto T^{i,j}_{(2p)} \propto T^{i,j}_{(3)} \propto U_s^2 \Delta\delta \Delta t = \beta U_s^3, \tag{8e}$$

where following Moser et al.[8]

$$\beta = \frac{1}{U_s} \frac{\Delta\delta}{\Delta t}. \tag{9}$$

Finally, the growth rate of the wake must be given by

$$\frac{\Delta\delta(t)^2}{\Delta t} = \kappa \rightarrow \delta(t) = \kappa^{1/2} (t - t^o)^{1/2} \tag{10}$$

in agreement with the results from the single-point analysis. Thus, the equations for the two-point velocity correlation tensor have self-similar solutions that include the self-similar solutions for the single-point moments[8] as a special case. This approach can also be extended to the governing equations for the multi-time correlations[6] by defining a transformed time scale given by

$$\tilde{t} = \ln(t/t_o),$$

and to equations for the higher-order and multi-point correlations. Thus, the hypothesis that the large-scale motions in this flow asymptotically evolve in an equilibrium manner is consistent with the governing equations of these measures of the structures.

There is also no unique choice for the similarity scales that will remove all of the flow dependent parameters from the equations. For example, if the constants of proportionality in equation 8 are taken as unity the momentum equations and the governing equations for $q_{\alpha\alpha}$ and $q_{1,2}$ become

$$-\frac{\Delta\eta f}{\Delta\eta} = -\frac{\Delta g}{\Delta\eta} \tag{11a}$$

$$\int_{-\infty}^{\infty} f(\eta)d\eta = -\gamma = -\frac{\dot{m}}{\rho}\frac{1}{U_s\delta} \tag{11b}$$

$$\begin{aligned}
-2q_{\alpha,\alpha} - \left\{\xi\frac{\partial}{\partial\xi} + \eta\frac{\partial}{\partial\eta} + \eta'\frac{\partial}{\partial\eta'} + \zeta\frac{\partial}{\partial\zeta}\right\} q_{\alpha,\alpha} + \left[\frac{1}{\kappa_1\beta}\right]\{f(\eta) - f(\eta')\}\frac{\partial q_{\alpha,\alpha}}{\partial\xi} = \\
-\left\{\frac{\delta_{\alpha1}}{\kappa_1}\frac{\partial}{\partial\xi}\left(\pi^{(1)}_{,1} - \pi^{(2)}_{1,}\right) + \left(\frac{\partial\pi^{(1)}_{,2}}{\partial\eta} + \frac{\pi^{(2)}_{2,}}{\partial\eta'}\right) + \frac{\delta_{\alpha3}}{\kappa_3}\frac{\partial}{\partial\zeta}\left(\pi^{(1)}_{,3} - \pi^{(2)}_{,3}\right)\right\} \\
-\left[\frac{\delta_{\alpha2}}{\beta}\right]\left\{\frac{1}{\kappa_1}\frac{\partial tt^{(1)}_{\alpha,\alpha}}{\partial\xi} + \frac{\partial tt^{(2)}_{\alpha,\alpha}}{\partial\eta} + \frac{\partial tt^{(2p)}_{\alpha,\alpha}}{\partial\eta'} + \frac{1}{\kappa_3}\frac{\partial tt^{(3)}_{\alpha,\alpha}}{\partial\zeta}\right\} - \left\{q_{2,1}\frac{df}{d\eta} + q_{1,2}\frac{df}{d\eta'}\right\}\delta_{\alpha1} \\
+\left[\frac{1}{Re_\delta\beta}\right]\left\{\frac{2}{\kappa_1^2}\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2} + \frac{\partial^2}{\partial\eta'^2} + \frac{2}{\kappa_3^2}\frac{\partial^2}{\partial\zeta^2}\right\} q_{\alpha,\alpha}\ ,
\end{aligned} \tag{11c}$$

$$\begin{aligned}
-2q_{1,2} - \left\{\xi\frac{\partial}{\partial\xi} + \eta\frac{\partial}{\partial\eta} + \eta'\frac{\partial}{\partial\eta'} + \zeta\frac{\partial}{\partial\zeta}\right\} q_{1,2} + \left[\frac{1}{\kappa_1\beta}\right]\{f(\eta) - f(\eta')\}\frac{\partial q_{\alpha,\alpha}}{\partial\xi} = \\
-\left[\frac{1}{\kappa_1\beta}\right]\frac{\partial\pi^{(1)}_{,2}}{\partial\xi} + \frac{\partial\pi^{(2)}_{1,}}{\partial\eta'} - \left[\frac{1}{\beta}\right]\left\{\frac{1}{\kappa_1}\frac{\partial tt^{(1)}_{1,2}}{\partial\xi} + \frac{\partial tt^{(2)}_{1,2}}{\partial\eta} + \frac{\partial tt^{(2p)}_{1,2}}{\partial\eta'} + \frac{1}{\kappa_3}\frac{\partial tt^{(3)}_{1,2}}{\partial\zeta}\right\} \\
-\left[\frac{1}{\beta}\right]q_{2,2}\frac{df}{d\eta} + \left[\frac{1}{Re_\delta\beta}\right]\left\{\frac{2}{\kappa_1^2}\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2} + \frac{\partial^2}{\partial\eta'^2} + \frac{2}{\kappa_3^2}\frac{\partial^2}{\partial\zeta^2}\right\} q_{1,2}\ ,
\end{aligned} \tag{11d}$$

where $\kappa_1 = \delta_1/\delta$ and $\kappa_3 = \delta_3/\delta$. The scales for the two-point velocity correlations could be chosen differently but there is no choice that will remove the parameter β from the equation. In particular, there is not way to remove the parameter β from both the second and third terms in the two-point equations (the unsteady term and energy transfer from the mean shearing and rotation). Thus, there is a whole family of possible 'equilibrium' solutions for the large-scale structures

in the wake, in agreement of the analysis of the single-point moments[5,8]. The equations and the experiments do not given any indication that there is an preferred or unique value of β for all wakes. Instead, they suggest that different 'equilibrium' structures could sustain different growth rates in the far field of the wake. In this case, the 'equilibrium' structures that occur in the flow would depend on the initial conditions of the wake as proposed by George[4].

This is contrary to the traditional conjecture that the temporally evolving wake should asymptotically evolve to an 'universal' solution characterized by only the mass deficit of the virtual source that generated the self-similar wake and the fluid properties. Thus, for a given Reynolds number the wake would evolve to a 'universal' solution characterized by the length and velocity scales given by

$$U_s = (\dot{m}/\rho)^{1/2} t^{-1/2} \tag{12a}$$

$$\delta = (\dot{m}/\rho)^{1/2} t^{1/2}. \tag{12b}$$

Goerge [4] argued this would not hold for flows from finite sources because it would require a number of length and velocity scales to characterize the initial conditions of these flows. This explanation seems incorrect for two reasons. First, the self-similar solutions are only formally valid for flows generated from virtual sources so if the flows from finite sources asymptotically approach self-similar solutions they must asymptotically approach the properties of these solutions. Second, the analysis of the governing equations shows that there is not *even* an unique self-similar solution for flows generated from a virtual sources.

The error in the traditional argument is that it implicitly assumes that the flow is generated from a steady virtual source that can be uniquely characterized by its mass deficit. However, the flow generated from a steady virtual source would have to undergo transition at some point to become turbulent and thus could not be self-similar over its entire evolution as required. Thus, self-similar turbulent flows must be generated from *unsteady* sources. It would require essentially an infinite number of invariants to uniquely characterize their unsteady nature so it is not unexpected that there ate families of unsteady sources that can produce self-similar flows. In fact, more surprisingly it seems that that unsteady sources that can produce self-similar flows can be characterized using only the parameter β and the Reynolds number.

2.1 Kolmogrov's Solutions for the Small-Scale Motions

The self-similar solutions outlined in the previous section describe the evolution of all the scales of motion in this constant-Reynolds-number flow. The form of the solution will change with turbulent Reynolds numbers and should be consistent with Kolmogorov's solutions in the limit of infinite Reynolds number if both solutions describe turbulent flows.

It can be shown that the two-point velocity gradient tensor in this flow can be written as

$$\overline{\frac{\partial u_i}{\partial x_k}\frac{\partial u'_j}{\partial x'_l}} = \left[\frac{Q^{i,j}}{\delta^2}\right] g_{i,j;k,l}(\xi, \eta, \eta', \zeta, *), \tag{13}$$

where $g_{i,j;k,l}$ is a self-similar solution. Thus, the dissipation can be written as

$$\epsilon = \lim_{x_k \to x'_k} 2\nu\overline{s_{ij}s'_{ij}} = \left[\frac{\nu U_s^2}{\delta^2}\right] d_l(\eta, *) = \left[\frac{U_s^3}{\delta}\right] d_h(\eta, *) \ , \tag{14}$$

where the s_{ij} is the fluctuating rate-of-strain tensor. Thus, the self-similar solutions for the wake permit both the traditional low- and high-Reynolds-number scaling for the dissipation.

It follows that the Kolmogorov length and velocity scale are proportional to the length and velocity scale for the large-scale motions; i.e.,

$$\eta_k = (\nu^3/\epsilon)^{1/4} \propto (\delta/U_s^3)^{1/4} \propto \delta \tag{15a}$$

$$u_k = (\nu\epsilon)^{1/2} \propto (U_s^3/\delta) \propto U_s \ , \tag{15b}$$

since $U_s \propto 1/\delta$. Thus, the similarity solution for the two-point velocity correlation tensor can be written in terms of the Kolmogorov variables. The analysis outlined here does not impose any restrictions on the functional form of the self-similar solutions so the description of the small-scale motions could asymptotically approach Kolomogorov's solutions in the infinite-Reynolds-number limit.

2.2 Comparison with Data

Although the analysis demonstrated that the equations have self-similar solutions, it is necessary to compare the predictions with data from real flows to determine if the solutions describe the evolution of these flows. Ewing[6] carried out an extensive comparison using data from two Direct Numerical Simulations of the temporally evolving wake[10,8] and found good agreement between the prediction and the data from the flows. For example, the one-dimensional spectra of the streamwise velocity component in the mean flows and lateral direction in similarity coordinates; i.e.,

$$\tilde{F}_{11}^1(\tilde{k}_1, \eta) = \int_{-\infty}^{\infty} q_{1,1}(\xi, \eta, \eta' = \eta, \zeta = 0) e^{\mathrm{i}\tilde{k}_1\xi} d\xi, \tag{16a}$$

$$\tilde{F}_{11}^3(\tilde{k}_3, \eta) = \int_{-\infty}^{\infty} q_{1,1}(\xi = 0, \eta, \eta' = \eta, \zeta) e^{\mathrm{i}\tilde{k}_3\zeta} d\zeta \ , \tag{16b}$$

at the centerline and half-deficit point in the wake are shown in figure 2. Here, U_s has been set equal to the centerline velocity deficit in the wake and δ the distance between half-velocity points. It is clear that the spectra in both directions collapse over the entire range of wavenumbers as predicted in theory.

3 Analogy Between the Temporally Evolving and the Self-Propelled Wake

The geometry of the temporally evolving wake resembles both a wake and a jet so it is not clear at first glance whether the evolution of the temporally evolving

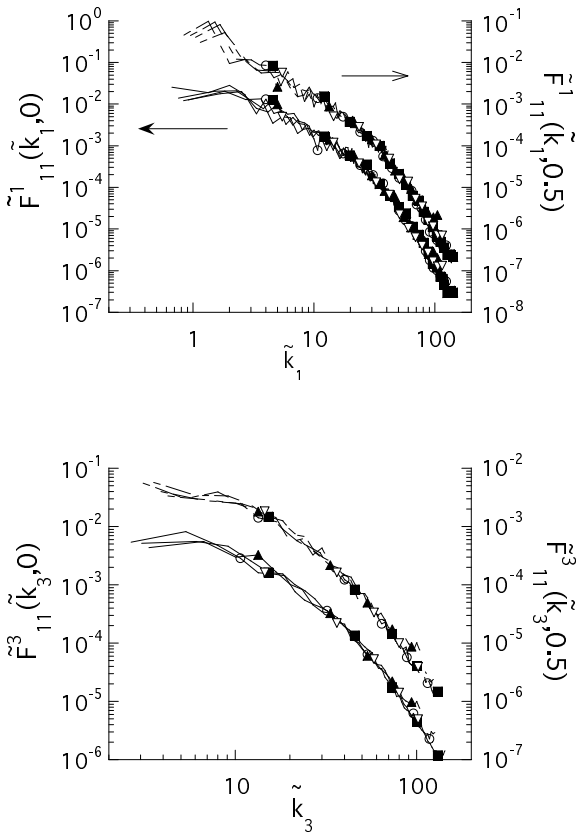

Fig. 2. Scaled one-dimensional spectra in (a) the mean flow direction, $\tilde{F}^1_{11}$ and (b) the lateral direction, $\tilde{F}^3_{11}$ at the centerline and half-deficit point in the forced wake simulation[8]. The spectra are taken from points where δ/δ_d = ——— 5.306, - - - - 6.636, —·— 6.636, and ——— 7.235 respectively.

wake is asymptotically analogous to a jet, a wake, or both. In order to show two flows are analogous it is necessary to show that the first order governing equations from one flow can be transformed to yield the governing equations in the other flow. An example is grid turbulence in a wind tunnel and decaying isotropic turbulence. It can be shown that the similarity equations for the single-point moments in the temporally evolving wake can be transformed to yield the first-order equations for the self-propelled wake. The equations for the two-point velocity correlation are examined here.

It is convenient to consider the wake generated by a flow past a stationary body. The equations for this flow can then be transformed to the self-propelled

wake by setting $t = \tilde{x}_1/U_\infty$ In the spatially evolving wake the governing equations for the two-point velocity correlations[9] can be reduced to

$$U_\infty \frac{\partial \overline{u_i u_j'}}{\partial \tilde{x}_1} + \underbrace{(\delta U_1 - \delta U_1')\frac{\partial \overline{u_i u_j'}}{\partial r_1}} + \underline{\frac{1}{2}(\delta U_1 + \delta U_1')\frac{\partial \overline{u_i u_j'}}{\partial \tilde{x}_1}} + \underline{U_2 \frac{\partial \overline{u_i u_j'}}{\partial x_2} + U_2' \frac{\partial \overline{u_i u_j'}}{\partial x_2'}}$$

$$-\frac{1}{\rho}\left[\frac{\partial}{\partial r_1}\left(\overline{pu_j'}\delta_{i1} - \overline{p'u_i}\delta_{j1}\right) + \underline{\frac{1}{2}\frac{\partial}{\partial \tilde{x}_1}\left(\overline{pu_j'}\delta_{i1} + \overline{p'u_i}\delta_{j1}\right)} + \frac{\partial \overline{pu_j'}}{\partial x_2}\delta_{i2} + \frac{\partial \overline{p'u_i}}{\partial x_2'}\delta_{j2} + \right.$$

$$\left. + \frac{\partial}{\partial r_3}\left(\overline{pu_j'}\delta_{i3} - \overline{p'u_i}\delta_{j3}\right)\right] - \frac{\partial}{\partial r_1}\overline{(u_1 - u_1')u_i u_j'} - \underline{\frac{1}{2}\frac{\partial}{\partial \tilde{x}_1}\overline{(u_1 + u_1')u_i u_j'}}$$

$$-\frac{\partial \overline{u_2 u_i u_j'}}{\partial x_2} - \frac{\partial \overline{u_2' u_i u_j'}}{\partial x_2'} - \frac{\partial}{\partial r_3}\overline{(u_3 - u_3')u_i u_j'} - \underbrace{\overline{u_j' u_2}\frac{\partial U_1}{\partial x_2}\delta_{i1}} - \underbrace{\overline{u_i u_2'}\frac{\partial U_1'}{\partial x_2'}\delta_{j1}}$$

$$+ \nu\left[2\frac{\partial^2}{\partial r_1^2} + \underbrace{\frac{1}{2}\frac{\partial^2}{\partial \tilde{x}_1^2}} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial {x_2'}^2} + 2\frac{\partial^2}{\partial r_3^2}\right]\overline{u_i u_j'} \tag{17}$$

where $\tilde{x}_1 = (x_1 + x_1')/2$ and $r_1 = x_1 - x_1'$. The terms highlighted with underbraces and underlines are the terms that either do not have an analogous term or differ from the terms in the equations for the temporally evolving wake. It is straightforward to show that the underlined terms are negligible compared to the free-stream convection term in the small-deficit limit. Similarly the highlighted viscous term is also higher order in reasonably high-Reynolds-number flows.

Using the self-similar solution for the mean velocity, it follows the mean velocity difference in the second term in this equation can be written as

$$\delta U_1 - \delta U_1' = U_s(x_1) f(\eta) - U_s(x_1') f(\eta'). \tag{18}$$

The scale $U_s(x_1)$ can be approximated as

$$U_s(x_1) = U_s(\tilde{x}_1)\left[1 - \frac{r_1}{4\tilde{x}_1} + \ldots\right]. \tag{19}$$

and a similar approximation for $U_s(x_1')$ it follows that

$$\delta U_1 - \delta U_1' = U_s(\tilde{x}_1)\left\{f(\eta) - f(\eta') - \frac{r_1}{2\tilde{x}_1}\left[f(\eta) + f(\eta')\right] + \ldots\right\}. \tag{20}$$

The characteristic length-scale of the structure in the streamwise direction grow at the same rate at the width of the wake so that

$$|r_1| \sim \delta(\tilde{x}_1) \sim (\tilde{x}_1 - \tilde{x}_1^o)^{1/2}.$$

Thus, the higher-order terms in the expansion become negligible as the flow evolves downstream. Physically, this implies that the straining and rotation

caused by the mean shear across the spatially evolving wake (also present in the temporally evolving wake) become much more important than the mean strain and rotation caused by the changes in the velocity deficit in the mean-flow direction. Similarly, it can be shown that the mean velocity gradients in the production terms can be evaluated at $\tilde{x}_1$ to first order. Note, in a flow that grows linearly, such as the plane jet exiting into a still ambient, the additional terms in the expansion are not necessarily higher order.

It can then be shown that the resulting first order equation for the two-point velocity correlations have self-similar solutions of the form[11]

$$\overline{u_i u'_j} = Q^{i,j}_{(s)}(\tilde{x}_1) q^{(s)}_{i,j}(\xi_s, \eta_s, \eta'_s, \zeta_s, *) \,, \tag{21a}$$

$$\overline{p u'_j} = \Pi^{j}_{(s1)}(\tilde{x}_1) \pi^{(s1)}_{,j}(\xi_s, \eta_s, \eta'_s, \zeta_s, *) \,, \tag{21b}$$

$$\overline{p' u_i} = \Pi^{i,}_{(s2)}(\tilde{x}_1) \pi^{(s1)}_{i,}(\xi_s, \eta_s, \eta'_s, \zeta_s, *) \,, \tag{21c}$$

$$\overline{(u_\beta - u'_\beta) u_i u'_j} = T^{i,j}_{(s\beta)}(\tilde{x}_1) t^{(s1)}_{i,j}(\xi_s, \eta_s, \eta'_s, \zeta_s, *) \,, \tag{21d}$$

$$\overline{u_2 u_i u'_j} = T^{i,j}_{(s2)}(\tilde{x}_1) t^{(s2)}_{i,j}(\xi_s, \eta_s, \eta'_s, \zeta_s, *) \,, \tag{21e}$$

$$\overline{u_i u'_2 u'_j} = T^{i,j}_{(s2p)}(\tilde{x}_1) t^{(s2p)}_{i,j}(\xi_s, \eta_s, \eta'_s, \zeta_s) \tag{21f}$$

where $\xi_s = r_1/\delta_1(\tilde{x}_1)$, $\eta_s = x_2/\delta(\tilde{x}_1)$, $\eta'_s = x'_2/\delta(\tilde{x}_1)$, and $\zeta_s = r_3/\delta_3(\tilde{x}_1)$ are the similarity variables analogous to those defined for the temporally evolving wake. It can also be shown that the governing equations for the similarity solutions in the spatially evolving wake can be transformed to yield the governing for the self-similar for the temporally evolving wake by setting $\tilde{x}_1 = U_\infty t$ [11]. Thus, the evolution of this measure of the large-scale structures in the temporally evolving wake are asymptotically analogous to the those in the self-propelled wake. The evolution of the two flows differ in the near field so temporally and self-propelled wakes started with analogous initial conditions will likely not yield the same flow in the far field.

4 Summary and Concluding Remarks

The analysis outline here has shown the hypotheses:

- The large-scale motions in the far field of many turbulent free shear flows evolve in an equilibrium manner.
- There is not an single 'universal' or unique set of structures that exist in the far field of these flows.

are consistent with the governing equations for the temporally evolving wake. Ewing [6] has shown the same for a number of other free-shear flows, including the plane and round jet. These hypotheses, which are supported by increasing amount of experimental evidence, are contrary to the tradition thinking on turbulent shear flows and will have a profound affect on the future research on turbulent free-shear flows and the control of these flows.

References

1. A.A. Townsend: *The Structure of Turbulent Shear Flow*, 1st edn. (Cambridge, London, 1956).
2. R. Narasimha: 'The utilities and drawbacks of traditional approaches'. In: *Whither Turbulence? Turbulence at the Crossroads*. ed. by J Lumley (Springer, Berlin, Heidelberg 1990) pp. 13–49.
3. I. Wygnanski, F. Champagne, and B. Marasli: J. Fluid Mech., **168**, 31 (1986).
4. W.K. George: 'The self-preservation of turbulent flows and its relation to initial conditions and coherent structures'. In: *Advances in Turbulence*. ed. by W.K. George and R.E. Arndt (Hemisphere, New York 1989) pp. 39–73.
5. W.K. George: 'Some New Ideas for Similarity of Turbulent Shear Flows'. In: *Proceedings of International Symposium of Turbulence Heat and Mass Transfer ed. by K. Hanjelic* (Elsevier, Amsterdam 1995).
6. D. Ewing. On multi-point similarity solutions in turbulent free shear flows. Ph. D. Thesis State University of New York at Buffalo, New York (1995).
7. P.R. Spalart: J. Fluid Mech. **187**, 67 (1988).
8. R.D. Moser, M.M. Rogers, and D. Ewing, J. Fluid Mech., **367**, 255 (1998)
9. J.O. Hinze: *Turbulence*. (McGraw-Hill, New York 1975)
10. R.D. Moser and M.M. Rogers: 'Direct simulation of a self-similar plane wake. NASA Technical Memorandum–TM 108815 (1994)
11. D. Ewing: 'Analogy Between Spatially and Temporally Evolving Wakes. Canadian Applied Mechanics Confonference Hamilton, ON (1999)

Quantum Kinetic Model of Turbulence

Amador Muriel

Data Transport Systems, 347 East 62nd St., New York, NY 10021 USA

Abstract. We study a system of two-level atoms which are allowed collision-induced radiative transitions between the two states. A linear, but non-local master equation describing the model is solved exactly, illustrating two results : the occurrence of average velocity reversal before it goes to zero, and the development of two counter-streaming motion from an initial one stream, both phenomena induced by radiation from the excited state. The experimental implications of this model on the onset of turbulence is discussed, and we review experimental results to date.

1 Motivation

One of the mysteries of turbulence is how an element of volume in a fluid could suddenly reverse its direction once a critical velocity is reached. If the laws of mechanics remain the same for all ranges of velocity, this behavior is difficult to understand, a problem expressed as the infinite growth of the time derivative of velocity in a fluid about to become turbulent. This approach is consistent with the prevailing view that turbulence is strictly a flow phenomenon to be explained by as yet undiscovered solutions of the Navier-Stokes equation. By contrast, the view adopted in this paper is that when a fluid becomes turbulent, it not only enters a new mathematical regime, but a new physical regime as well. To get to the point, we propose that identifiable new internal degrees of freedom are excitable, so that fluid flow must account for the exchange of energy between the internal degrees of freedom, and the usual translational degrees of freedom. Where do these internal degrees of freedom come from? They could occur because of molecular clustering, or just plain molecular rotations. This view immediately takes on a molecular component that is not considered essential in the conventional continuum theory of hydrodynamics. For if the reality of molecules were accepted, a straightforward continuum application of the Navier-Stokes equation may not be sufficient to explain the onset of turbulence. We must then look at the threshold energy needed to excite the internal degrees of freedom. In the case of molecular rotations, the energy involved is of the order of kT/40 for diatomic molecules, well within reach of mechanical or thermal agitation. Once it is accepted that a threshold energy is needed to excite most of the molecules, a physical interpretation of the critical Reynolds number becomes possible : when energy is pumped into a system, exceeding the energy density needed to excite, and by the way, de-excite most of the molecules, turbulence begins. When such a regime is reached, probability density distributions for at least two species of the

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 103–113, 2001.
© Springer-Verlag Berlin Heidelberg 2001

molecule are needed (excited and un-excited). If transport equations analogous to the Navier-Stokes equation are to be derived theoretically, there would at least be two such equations, and all macrosocopically averaged quantities would contain at least two contributions. Furthermore, the equations will be coupled – it is such coupling from which we find turbulence developing.

Hence, we propose the following picture : in low energy systems, a single Navier-Stokes equation is sufficient, a continuum model could neglect the unexcited internal degrees of freedom. When more energy is pumped into the system, the internal degrees of freedom exchange energy with translational degrees of freedom. The equations describing such an energetic system – no longer a single Navier-Stokes equation – admit turbulent-like solutions. To show these features of a molecular point-of-view, we create a new model that displays both velocity reversal and turbulent-like characteristics. Since we will need to describe multi-species at the outset, ruling out the use of only one Navier-Stokes equation, we concentrate first on the more fundamental kinetic theory from which transport equations may be derived much later.

2 Kinetic Equations

The kinetic theory of gases with molecules possessing internal degrees of freedom is not so well-developed because the Boltzmann collision expression is meant to describe elastic collisions. In a gas of two-level atoms, for example, there must be one distribution function for each specie, and the Boltzmann collision term must describe collisions between the molecules of the same specie, or state, as well as a collision term between the two species, or states. The Boltzmann description will introduce two coupled transport equations, each of which is already non-linear. Given that the conventional one-specie Boltzmann transport equation is unsolved, such a system of coupled, non-linear equations will be even more difficult to solve. Nevertheless, there is a need to study the kinetic theory of molecules with an internal degree of freedom that could result in radiation. We want to see if any new qualitative change in the description of such a gas could arise.

In this work , we avoid the traditional Boltzmann collision terms and instead use measurable transition probabilities between quantum states. To simplify the problem even more, we eventually consider only a gas of two-level atoms in one dimension. The decay of the excited state will result in the emission of radiation. So would collisional excitation and de-excitation. We assume that radiation ensues. In this paper, such radiative transitions, by natural decay, or collisional excitation, will be assumed to radiate the same frequency of light. The radiation produced is assumed to participate no more in the evolution of the system, resulting in dissipation. Our approach in this work differs considerably from realistic but incomplete studies [1,2,3]. But our simple approach results in an exactly solvable case, a novelty, allowing us to make plausibility statements about a subject as complex as turbulence.

To illustrate, we may write two coupled master equations describing the distribution functions for each of the two species. As there are very few exactly solvable master equations [4] , any new addition to the collection of exactly solvable master equation is also interesting.

Following [4], let the two internal states of a molecule be separated by the energy gap Δ with explicitly defined transitions. In each of the two states, separated by the gap energy , the single particle distribution function at time is given by $f_2(r,p,t)$ and $f_1(r,p,t)$. Following the master equation approach in non-equilibrium statistical mechanics [3], we write the following equations for the distribution functions :

$$\frac{\partial f_2}{\partial t} + \frac{p}{m}\frac{\partial f_2}{\partial r} = -\gamma_{21} f_2 + \gamma_{12}\int dp' f_1(r,p',t)\delta(\frac{p'^2}{2m} - \frac{p^2}{2m} - \Delta) \tag{1}$$

$$\frac{\partial f_1}{\partial t} + \frac{p}{m}\frac{\partial f_1}{\partial r} = -\gamma_{12} f_1 + \gamma_{21}\int dp' f_2(r,p',t)\delta(\frac{p'^2}{2m} - \frac{p^2}{2m} - \Delta) \tag{2}$$

The left hand sides of Eqs. (1,2) are the usual streaming terms. What would otherwise be Boltzmann collision terms are replaced by probability transition rates.

The first term on the rhs of Eq. (1) represents the change in the probability distribution due to decay to the ground state, the second term is the increase in the probability by scattering from the ground state, of a particle with initial energy $p^2/2\theta m + \Delta$, ending up with kinetic energy $p^2/2m$ in the excited state. The energy Δ is given up as unrecoverable radiation.

The first term in Eq. (2) contains the transition probability γ_{21} , which includes both the natural decay constant and the collision induced transition. The second term is the loss of probability due to collision induced transition to the excited state with transition rate γ_{12} , giving up radiation with energy Δ .

We use the delta function to conserve total energy. In the above model equations, momentum is not conserved locally due to the semi-classical radiation.

Specifically excluded from our model kinetic equations are a priori non-linear products of any two distribution functions. It would appear that nothing interesting could arise from this simplification. To our surprise, enough interesting behavior remains to justify studying our model, allowing us to ask the question whether complex behavior is the province of non-linearities alone. From a mathematical point of view, as we will see later, our non-local expressions seem to simulate a weak form of non-linearity, which, however, produces time dependent distribution functions of some complexity.

We have written Eqs. (1) and (2) to provide a physical justification for our model, but we may abstract the equations even more when written in the following form :

$$\frac{\partial}{\partial t}\begin{pmatrix} f_2 \\ f_1 \end{pmatrix} = \begin{pmatrix} D - \gamma_{21} & \gamma_{12} J \\ \gamma_{21} J & D - \gamma_{12} \end{pmatrix}\begin{pmatrix} f_2 \\ f_1 \end{pmatrix} \tag{3}$$

where we have put the displacement operator $D = -\frac{p}{m}\frac{\partial}{\partial r}$ and J , a jump operator, yet undefined until later.

Were the matrix elements only numbers, the solution of the above linear equation is trivial, but because of the presence of non-commuting operators, the exact solution of Eq. (3) needs to be derived with some care. Although linear, Eq. (3) could result in functions of displaced coordinates and momenta, curiously making the equation non-local in momentum. The non-linearity of the Boltzmann equation is replaced by non-locality, perhaps a weak form of non-linear behavior. At this point however, J is no more than a mathematical operator and may be redefined to describe a model independent of our motivation.

3 Solution of Model Equation

We may define the matrix components

$$\begin{aligned} L_{11} &= D - \gamma_{21} \\ L_{12} &= \gamma_{12} J \\ L_{21} &= \gamma_{21} J \\ L_{22} &= D - \gamma_{12} \end{aligned} \tag{4}$$

which then allows us to write a formal solution [5] of Eq. (3)

$$\begin{aligned} f_2 = & \left\{ e^{\left[(L_{11}+L_{22})+\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}} + e^{\left[(L_{11}+L_{22})-\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}} \right\} \\ & \frac{f_2(0)}{2} \\ - & \left\{ e^{\left[(L_{11}+L_{22})+\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}} - e^{\left[(L_{11}+L_{22})-\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}} \right\} \\ & \frac{(L_{11}-L_{22})}{\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}} \frac{f_2(0)}{2} \\ + & \left\{ e^{\left[(L_{11}+L_{22})+\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}} - e^{\left[(L_{11}+L_{22})-\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}} \right\} \\ & \frac{L_{12}}{\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}} f_1(0) \end{aligned} \tag{5}$$

$$\begin{aligned} f_1 = & \left\{ e^{\left[(L_{11}+L_{22})+\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}} + e^{\left[(L_{11}+L_{22})-\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}} \right\} \\ & \frac{f_1(0)}{2} \\ + & \left\{ e^{\left[(L_{11}+L_{22})+\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}} - e^{\left[(L_{11}+L_{22})-\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}} \right\} \\ & \frac{(L_{11}-L_{22})}{\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}} \frac{f_1(0)}{2} \end{aligned}$$

$$+\left\{e^{\left[(L_{11}+L_{22})+\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}}-e^{\left[(L_{11}+L_{22})-\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}\right]\frac{t}{2}}\right\}$$

$$\frac{L_{21}}{\sqrt{(L_{11}-L_{22})^2+4L_{12}L_{21}}}f_2(0) \tag{6}$$

The method used to arrive at the above solution is shown elsewhere [5] which serves to correct and extend an earlier paper. The biggest problem in the above formal solution is the non-commutativity of the displacement operator and the jump operator . In this paper, we extend our earlier work [5] to generalize to a non-uniform system as promised in that paper.

Let us put

$$\Gamma = \gamma_{21}+\gamma_{12},\ \ \gamma = \gamma_{21}-\gamma_{12},\quad \gamma_o^2 = \gamma_{12}\gamma_{21} \tag{7}$$

then

$$L_{11}+L_{22} = 2D-\Gamma \tag{8}$$

$$(L_{11}-L_{22})^2+4L_{12}L_{21} = \gamma^2+4\gamma_o^2J^2 \tag{9}$$

The key to the simplification of the formal solution containing non-commuting operators D and J is to expand the exponential operators. With the use of some operator identities as explained in Reference [5], we get, in expanded operator form

$$\begin{aligned} f_2 = e^{-\Gamma t/2}e^{-Dt}\Bigg\{&\sum_{m=0}^{\infty}\frac{(\gamma^2+4\gamma_o^2J^2)^m(t/2)^{2m}f_2(0)}{(2m)!}\\ &-\frac{\gamma t}{2}\sum_{n=0}^{\infty}\frac{(\gamma^2+4\gamma_o^2J^2)^n(t/2)^{2n}f_2(0)}{(2n+1)!}\\ &+\frac{\gamma_{12}t}{2}\sum_{n=0}^{\infty}\frac{(\gamma^2+4\gamma_o^2J^2)^n(t/2)^{2n}Jf_1(0)}{(2n+1)!}\Bigg\} \end{aligned} \tag{10}$$

$$\begin{aligned} f_1 = e^{-\Gamma t/2}e^{-Dt}\Bigg\{&\sum_{m=0}^{\infty}\frac{(\gamma^2+4\gamma_o^2J^2)^m(t/2)^{2m}f_1(0)}{(2m)!}\\ &+\frac{\gamma t}{2}\sum_{n=0}^{\infty}\frac{(\gamma^2+4\gamma_o^2J^2)^n(t/2)^{2n}f_1(0)}{(2n+1)}\\ &+\frac{\gamma_{21}t}{2}\sum_{n=0}^{\infty}\frac{(\gamma^2+4\gamma_o^2J^2)^n(t/2)^{2n}Jf_2(0)}{(2n+1)!}\Bigg\} \end{aligned} \tag{11}$$

Eqs. (10) and (11) render the formulas ready for numerical implementation provided the jump operator J is properly defined. It is only at this point where we introduce the one dimensionality of the results. We have yet to use Eqs. (10,11) in non-uniforms systems, but for now, we render exact some conclusions arrived at in [5] with approximations.

4 Uniform System

We now choose the operator J by putting

$$Jf(p) = \frac{1}{2}\left[f(\sqrt{p^2+\delta}) + f(-\sqrt{p^2+\delta})\right] \tag{12}$$

which allows us to write

$$J^j f(p) = \frac{1}{2}\left[f(\sqrt{p^2+j\delta}) + f(-\sqrt{p^2+j\delta})\right] \tag{13}$$

smallskip for $j > 0$. Eq. (13) may be checked by repeated application of J . We have put $\delta = 2m\Delta$.

Suppose that we start with

$$f_2(0) = \sqrt{\frac{\beta}{\pi}} e^{-\beta(p-p_o)^2} \quad , \quad f_1(0) = 0 \tag{14}$$

then

$$f_2 = e^{-\Gamma t/2} e^{-Dt} \left\{ \sum_{m=0}^{\infty} \frac{(\gamma^2 + 4\gamma_o^2 J^2)^m (t/2)^{2m} f_2(0)}{(2m)!} - \frac{\gamma t}{2} \sum_{n=0}^{\infty} \frac{(\gamma^2 + 4\gamma_o^2 J^2)^n (t/2)^{2n} f_2(0)}{(2n+1)!} \right\} \tag{15}$$

$$f_1 = e^{-\Gamma t/2} e^{-Dt} \left\{ \frac{\gamma_{21} t}{2} \sum_{n=0}^{\infty} \frac{(\gamma^2 + 4\gamma_o^2 J^2)^n (t/2)^{2n} J f_2(0)}{(2n+1)!} \right\} \tag{16}$$

By using the sum of the the two probabilities to calculate the average momentum, we get

$$< p >= p_o e^{-(\gamma_{12}+\gamma_{21})t/2} \left[1 - (\gamma_{21} - \gamma_{12})t/2\right] \tag{17}$$

which will produce momentum reversal when more commonly expected magnitudes of the transition probabilities are used. In arriving at Eq. (17) it suffices to observe the parity of the term by term contributions to the average momentum.

Thus, when the excited state is first pumped, by electromagnetic, or even mechanical means, momentum reversal occurs when $\gamma_{21} > \gamma_{12}$.

We also find that when the all the particles are initially in the ground state, momentum reversal can only occur in the unlikely or uncommon case of a metastable excited state, or $\gamma_{21} < \gamma_{12}$.

It is also interesting to study the behavior of the system when the excited state is initially unpopulated, with all particles in the ground state with a displaced Maxwellian as above. There is no momentum reversal, but we observe the formation of two counter streams from an initial one stream. We illustrate in Fig. 1 a,b,c the approach to zero of the average momentum by displaying

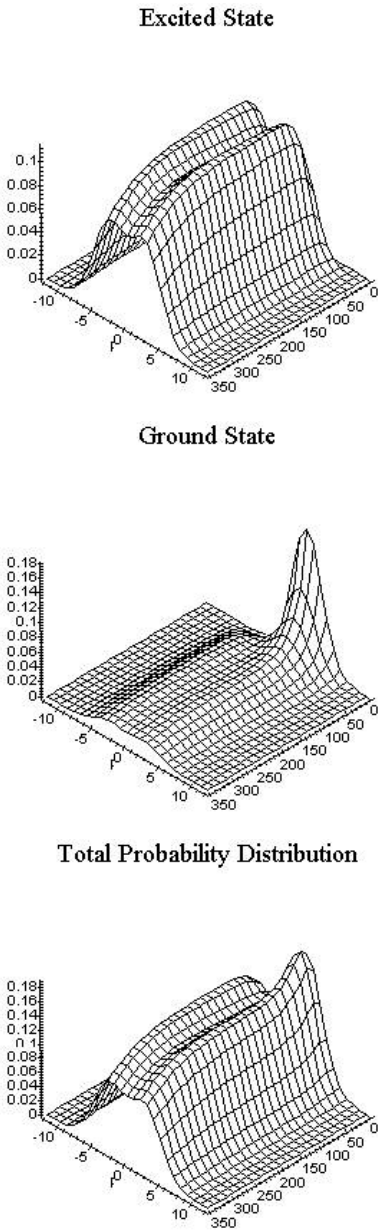

Fig. 1. Typical time evolution of the momentum distributions for the excited state (top), ground state (middle), and both states (bottom), showing the development of counter streaming momenta from an initial one stream initial condition in the ground state.

the time evolution of the two distribution functions and their sum using the following parameters : $\gamma_{21} = 0.05,\ \ \gamma_{12} = 0.04,\ \ \beta = 0.1,\ \ \delta = 0.2,\ \ p_o = 3$.

We start ith all particles in the ground state with Maxwellian distribution with mean momentum 3. Take 8 terms of the infinite series and follow the time evolution of the momentum distributions until time 350.

The excited state, initially empty, is gradually populated with two symmetric counter streaming terms, depleting the ground state by collisions. The net result is a decelerating main stream, modified by a smaller counterstream, as the excited state begins to depopulate. Our simple model generates the formation of counter streaming by radiation. Yet another possible interpretation is that we see projected in one dimension what may well be eddy formation in three dimensions, an observation that we hope to study when we go to three dimensions, a doable, but more ambitious study, outside the scope of this work.

It thus appears that once we allow radiation to be emitted isotropically, in this case equally probable in both directions, we might resolve two problems in turbulence : (1) how fluid motion could decelerate spontaneously and even reverse momentum, and (2) how one could develop eddies. As far as we can tell, the behavior that we illustrate here has heretofore not been seen analytically. The mysteries of spontaneous momentum reversal and eddy formation are not so unexplainable in our kinetic model. In a future paper, we will consider a system driven by external forces as well.

5 Conclusions from the Model

We have introduced an ad-hoc master equation to model a uniform system of two level atoms by writing down time evolution equations for the classical momentum distribution functions. We conserve energy, but not momentum, allowing radiation to carry both energy and momentum in both directions in our one dimensional model. We observe two things (1) the possibility of momentum reversal, as observed in turbulence, and the formation of two counterstreaming motion where we start with only one, again observed in nature. In our model, both phenomena are due to radiation from the excitation and de-excitation of internal states.

The equations we use are very simple, linear, but non-local in momentum. The master equations are exactly solvable, which is their main virtue, and allows us to find behavior reminiscent of real systems. The main feature of this model is the presence of internal states which allow for radiation, suggesting to us that the inclusion of internal molecular states significantly alters the expected flow regime intuitively expected to arise in such a simple flow. These results were already hinted in [5] by approximations, this work is exact. In addition, the present work extends the formal solution in [5] to a non-uniform system, in a form useful for subsequent studies, and stresses the development and evolution of the counterstreaming terms, an analytic result that may carry more significance than brute force numerical simulations.

6 Experimental Consequences of the Model

Our main theoretical conclusion is that a two-level system can produce velocity reversal, an exact rigorous result from the model equations. The key physical reason for this reversal is the assumption of isotropic radiation that conserves energy but not local momentum. This model assumes the existence of an energy gap in the energy spectrum which could be overcome by the kinetic energy of collisions. But as different molecules possess different energy spectra, the onset of validity of our model, or the onset of turbulence, is molecule-dependent. This means that the ratio of two critical Reynolds numbers for the same apparatus for two different gases cannot be unity, as conventionally accepted, not without much accurate testing in the past – as it turns out – for such experiments have been traditionally performed with only one fluid. We thus predict that each pure gas will have its own different critical Reynolds number as opposed to the conventioanl view. Figure 2 shows the latest experimental data obtained by S. Novopashin and co-workers [9,10] and confirmed by J. Sommeria [11], in contradiction to the early null results of White and Sreenivasan [12].

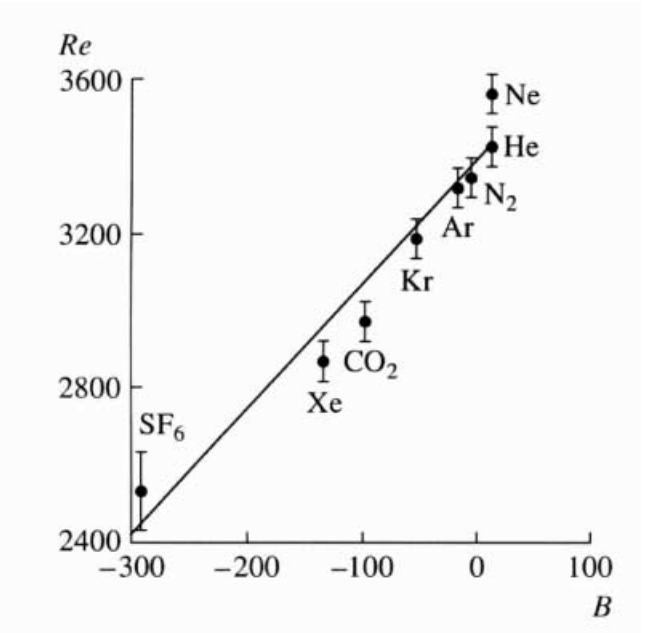

Fig. 2. Critical Reynolds numbers for different gases. According to conventional hydrodynamics, all points should lie in one horizontal line. The points are plotted against the second virial of the gases, highlighting the atomic and molecular differences. Courtesy of S. Novopashi.

There are other predictions of a molecular theory which this paper illustrates. These are the following : any physical process that excites the higher energy states of the molecules will prematurely excite turbulence, like sound waves [11] or electromagnetic waves, which may be called "non-inertial" contributions to turbulence, adopting the terminology of Groisman and Steinberg [13]. This will result in lower critical Reynolds numbers, already observed with polymeric molecules, whose internal degrees of freedom, by an extension of our ideas, are even much easier to excite [13].

Finally, although our results in this paper are strictly theoretical, we must make comparisons with experiments because of the fundamental change in approach that we advocate. The suggestion that there must exist excitable atomic or molecular degrees of freedom before one gets into a regime of turbulence is quite drastic and needs further justification. For example, one must discuss the question of what constitutes an internal degree of freedom for atoms of the noble gases. We find that quantum confinement, as in the cell model of fluids [14] is sufficient to generate quantum levels of energy which are excitable by atomic collisions. In fact we propose a general rule: find all possible quantum levels for atoms, molecules, or clusters, then choose the lowest energy of excitation and use such energy thresholds as the threshold for the onset of turbulence. For noble gases, the energy states of quantum confinement suffice, for diatomic molecules, the quantized rotational states are lowest in energy.This argument seems to work for many cases of the onset of turbulence, including patently obvious quantum systems like bose condensates which become turbulent. In such situations, quantum confinement energies are easily calculated from the densities and the masses of the atoms or molecules which form the condensate [15].

In short,using our theoretical model, we find that turbulence may be explained by atomic and molecular physics, and by extension, quantum concepts. In conclusion, we suggest that turbulence is a manifestation of molecular structure, traditionally ignored in hydrodynamics, and even more philosophically, that turbulence is a macroscopic, observable consequence of quantum physics.

The assistance of the ICSC World Laboratory in Lausanne is gratefully acknowledged.

*E-mail : dtsny@bellatlantic.net

References

1. E. B. Davies, Quantum Theory of Open Systems (Academic Press , London 1976)
2. R. Monaco and L. Preziosi, Fluid Dynamic Applications of the Discrete Boltmann Equation (World Scientific, New Jersey 1991)
3. Yu. L. Klimontovich, The Kinetic Theory of Electromagnetic Processes (Springer-Verlag, Berlin 1983)
4. N.G. van Kampen, Stochastic Processes in Physics and Chemistry (North Holland, Amsterdam 1992)
5. A. Muriel, Physica D **124** (1998) 225-247, A. Muriel, to be published.
6. L. Sirovich (Ed.), New Perspectives in Turbulence (Springer, New York 1991)
7. A. Muriel, Phys. Rev. Lett. **75** (1995) 2255.

8. A. Muriel, Phys. Rev. A **50** (1994) 4286.
9. S. Novopashin, and A. Muriel, Tech. Phys. Lett. **26** (2000) 18.
10. O. Nerushev and S. Novopashin, Phys. Lett. A **232** (1997) 243.
11. J. Sommeria, to be published.
12. C. White and K. Sreenivasan, Phys. Lett. A **238** (1998) 323.
13. A. Groisman and V. Steinberg, Nature **405** (2000) 53.
14. See for example, T. L. Hill, An introduction to statistical thermodynamics, Dover Publications, New York, 1960.
15. C. Raman, et. al.,Phys. Rev. Lett.**83**, No 13, 2502-2505 (1999).

The Transport of Small Particles by a Fluid

Armando Babiano[1], Julyan H.E. Cartwright[2], Oreste Piro[3], and Antonello Provenzale[4]

[1] Laboratoire de Météorologie Dynamique, École Normale Supérieure, F-75231 Paris, France
[2] Laboratorio de Estudios Cristalográficos, CSIC, E-18071 Granada, Spain
[3] Institut Mediterrani d'Estudis Avançats, IMEDEA (CSIC–UIB), E-07071 Palma de Mallorca, Spain
[4] Istituto di Cosmogeofisica, CNR, I-10133 Torino, Italy

Abstract. The study of trajectories of particles suspended in a fluid in motion is of relevance to many problems; as much those in applied technology as those of theoretical interest. Properties of emulsions, dispersion of contaminants in the atmosphere or the oceans, sedimentation, and mixing, are just a few applications. From the point of view of dynamical systems, this type of problem motivates the study of models with a dimension that is relatively low, but sufficient to encounter theoretical aspects that are as yet little understood.

1 Passive Scalars

In the simplest formulation, the dynamics of a passive scalar — a marked element of the fluid, or a particle sufficiently small so as not to appreciably perturb the flow, but sufficiently large not to be perturbed by Brownian motion — in an incompressible fluid, can be described by the dynamical system

$$\frac{d\boldsymbol{r}}{dt}(x,y,z,t) = \boldsymbol{u}(x,y,z,t), \tag{1}$$

$$\nabla \cdot \boldsymbol{u} \qquad = 0, \tag{2}$$

where $\boldsymbol{r}$ is the position of the particle and $\boldsymbol{u}$ is the velocity of the fluid.

In two spatial dimensions, such a system is Hamiltonian; the stream function Ψ generates a velocity field

$$\dot{x} = v_x = \frac{\partial \Psi}{\partial x}(x,y,t), \tag{3}$$

$$\dot{y} = v_y = -\frac{\partial \Psi}{\partial y}(x,y,t). \tag{4}$$

The phase space — which is real space in this case — of such Hamiltonian systems typically shows a mixture of chaotic and integrable trajectories. This phenomenon of deterministic chaos, as distinct from turbulence, in fluid flows is termed Lagrangian turbulence or chaotic advection ([18,9]). Three-dimensional flows, not being Hamiltonian systems ([6]), present novel forms of chaotic advection, including resonance-induced dispersion, present in a class of unsteady

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 114–124, 2001.
© Springer-Verlag Berlin Heidelberg 2001

three-dimensional flows, in which the barriers formed by integrable regions in the flow break down allowing a single particle trajectory access to the whole of the phase space (Cartwright, Feingold, and Piro (1995,1996,1999)).

In two-dimensional, and most three-dimensional flows, however, the integrable regions of KAM tori in these chaotic systems do represent barriers to the transport of passive scalars, and a perturbation of practical interest which can annul the effect of these barriers is the possession of inertia by the transported particles.

2 Inertial Particles: Massive Passive Scalars

The equation of motion for a small, spherical tracer in an incompressible fluid may be written as

$$\begin{aligned}\rho_p \frac{d\mathbf{v}}{dt} = {} & \rho_f \frac{D\mathbf{u}}{Dt} + (\rho_p - \rho_f)\mathbf{g} \\ & -\frac{9\nu\rho_f}{2a^2}\left(\mathbf{v} - \mathbf{u} - \frac{a^2}{6}\nabla^2\mathbf{u}\right) \\ & -\frac{\rho_f}{2}\left(\frac{d\mathbf{v}}{dt} - \frac{D}{Dt}\left[\mathbf{u} + \frac{a^2}{10}\nabla^2\mathbf{u}\right]\right) \\ & -\frac{9\rho_f}{2a}\sqrt{\frac{\nu}{\pi}}\int_0^t \frac{1}{\sqrt{t-\zeta}}\frac{d}{d\zeta}\left(\mathbf{v} - \mathbf{u} - \frac{a^2}{6}\nabla^2\mathbf{u}\right)d\zeta. \end{aligned} \tag{5}$$

$\mathbf{v}$ represents the velocity of the particle, $\mathbf{u}$ that of the fluid, ρ_p the density of the particle, ρ_f, the density of the fluid it displaces, ν, the kinematic viscosity of the fluid, a, the radius of the particle, and $\mathbf{g}$, gravity. The derivative $D\mathbf{u}/Dt$ is along the path of a fluid element

$$\frac{D\mathbf{u}}{Dt} = \frac{\partial\mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u}, \tag{6}$$

whereas the derivative $d\mathbf{u}/dt$ is taken along the trajectory of the particle

$$\frac{d\mathbf{u}}{dt} = \frac{\partial\mathbf{u}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{u}. \tag{7}$$

The terms on the right of Eq. (5) represent respectively the force exerted by the undisturbed flow on the particle, buoyancy, Stokes drag, the added mass due to the boundary layer of fluid moving with the particle, and the Basset–Boussinesq force ([5,4]) that depends on the history of the relative accelerations of particle and fluid. The terms in $a^2\nabla^2\mathbf{u}$ are the Faxén [13] corrections. Equation (5) is derived under the assumptions that the particle radius and its Reynolds number are small, as are the velocity gradients around the particle. The equation is as given by Maxey and Riley [15], except for the added mass term of Taylor [23] (see [1]).

The most general description of the dynamics of these particles presents an enormous richness of phenomena (see, for example, [11,22,25,16]). It is characteristic of the dynamics for small inertia that, in the time-independent case, what were invariant surfaces in the model without inertia are transformed into spirals, due to centrifugal forces. As a consequence, particles tend to accumulate at separatrices of the flow. Except in the proximity of fixed points, the relative velocity of particle and fluid tends to zero at long times. For large inertia, particles are no longer confined within vortices. Stokes drag is the most important force acting in this case, so to a first approximation one can discard the other terms in Eq. (5) to obtain

$$\frac{d^2x}{dt^2} = -\mu\left(\frac{dx}{dt} - u_x(x,y,t)\right), \tag{8}$$

$$\frac{d^2y}{dt^2} = -\mu\left(\frac{dy}{dt} - u_y(x,y,t)\right), \tag{9}$$

where we are considering a horizontal layer with $\mathbf{v} = (dx/dt, dy/dt)$, and $\mu = 9\nu\rho_f/(2a^2\rho_p)$. This is a highly-dissipative and singular perturbation of a Hamiltonian system, with a four-dimensional phase space:

$$\dot{x} = p_x, \tag{10}$$

$$\dot{p}_x = -\mu(p_x - u_x(x,y,t)), \tag{11}$$

$$\dot{y} = p_y, \tag{12}$$

$$\dot{p}_y = -\mu(p_y - u_x(x,y,t)). \tag{13}$$

In the time-dependent case, particles tend to accumulate in caustics, which correspond to the chaotic regions of the model without inertia, as the upper panel of Fig. 1 shows. The relative velocity fluctuates chaotically, due to macroscopic, nonturbulent fluctuations, that act to give the particles deterministic but Brownianlike motion, illustrated in the lower panel of Fig. 1.

3 Neutrally Buoyant Particles

Let us now consider whether even in the most favourable case of neutral buoyancy a finite sized tracer particle may always be faithful to a flow trajectory. With this in mind, we set $\rho_p = \rho_f$ in Eq. (5). At the same time we assume that it be sufficiently small that the Faxén corrections be negligible. Furthermore we exclude the Basset–Boussinesq term, for we wish to obtain a minimal model ([21,10]) with which we may perform a mathematical analysis of the problem. If in this model there appear differences between particle and flow trajectories, with the inclusion of further terms these discrepancies will remain or may even be enhanced. If we rescale space, time, and velocity by scale factors L, $T = L/U$, and U, we arrive at the expression

$$\frac{d\mathbf{v}}{dt} = \frac{D\mathbf{u}}{Dt} - \mathrm{S}t^{-1}\,(\mathbf{v} - \mathbf{u}) - \frac{1}{2}\left(\frac{d\mathbf{v}}{dt} - \frac{D\mathbf{u}}{Dt}\right), \tag{14}$$

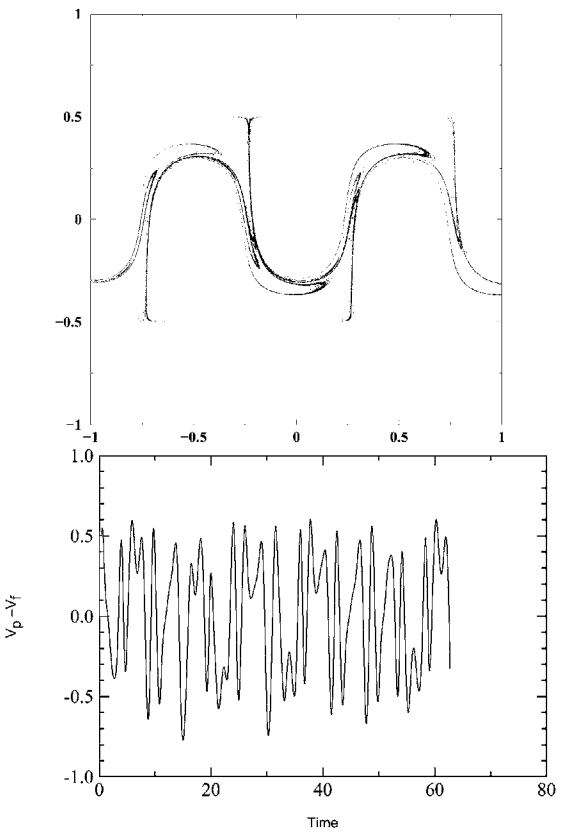

Fig. 1. Above: dense particles collect in fractal structures around a separatrix. Below: in the large inertia limit the relative velocity of particle and flow fluctuates chaotically, in a Brownianlike fashion.

where $\mathrm{S}t$ is the particle Stokes number $\mathrm{S}t = 2a^2U/(9\nu L) = 2/9\,(a/L)^2\mathrm{R}e_f$, $\mathrm{R}e_f$ being the fluid Reynolds number. The assumptions involved in deriving Eq. (5) require that $\mathrm{S}t \ll 1$.

In the past it has been assumed that neutrally buoyant particles have trivial dynamics (e.g., [10,11]), and the mathematical argument used to back this up is that if we make the approximation $D\mathbf{u}/Dt = d\mathbf{u}/dt$, which can be seen as a rescaling of the added mass, the problem becomes very simple

$$\frac{d}{dt}(\mathbf{v} - \mathbf{u}) = -\frac{2}{3}\,\mathrm{S}t^{-1}(\mathbf{v} - \mathbf{u}). \tag{15}$$

Thence

$$\mathbf{v} - \mathbf{u} = (\mathbf{v}_0 - \mathbf{u}_0)\exp(-2/3\,\mathrm{S}t^{-1}\,t), \tag{16}$$

from which we infer that even if we release the particle with a different initial velocity $\mathbf{v}_0$ to that of the fluid $\mathbf{u}_0$, after a transient phase the particle velocity

will match the fluid velocity, $\mathbf{v} = \mathbf{u}$, meaning that following this argument, a neutrally buoyant particle should be an ideal tracer. Although from the foregoing it would seem that neutrally buoyant particles represent a trivial limit to Eq. (5), this would be without taking into account the correct approach to the problem that has $D\mathbf{u}/Dt \neq d\mathbf{u}/dt$. If we substitute the expressions for the derivatives in Eqs. (6) and (7) into Eq. (14), we obtain

$$\frac{d}{dt}\left(\mathbf{v}-\mathbf{u}\right) = -\left(\left(\mathbf{v}-\mathbf{u}\right)\cdot\nabla\right)\mathbf{u} - \frac{2}{3}\,\mathrm{St}^{-1}\left(\mathbf{v}-\mathbf{u}\right). \tag{17}$$

We may then write $\mathbf{A} = \mathbf{v} - \mathbf{u}$, whence

$$\frac{d\mathbf{A}}{dt} = -\left(J + \frac{2}{3}\,\mathrm{St}^{-1} I\right)\cdot\mathbf{A}, \tag{18}$$

where J is the Jacobian matrix — we now concentrate on two-dimensional flows $\mathbf{u} = (u_x, u_y)$ —

$$J = \begin{pmatrix} \partial_x u_x & \partial_y u_x \\ \partial_x u_y & \partial_y u_y \end{pmatrix}. \tag{19}$$

If we diagonalize the matrix we obtain

$$\frac{d\mathbf{A}_D}{dt} = \begin{pmatrix} \lambda - 2/3\,\mathrm{St}^{-1} & 0 \\ 0 & -\lambda - 2/3\,\mathrm{St}^{-1} \end{pmatrix}\cdot\mathbf{A}_D, \tag{20}$$

so if $Re(\lambda) > 2/3\,\mathrm{St}^{-1}$, $\mathbf{A}_D$ may grow exponentially. Now λ satisfies $\det(J-\lambda I) = 0$, so $\lambda^2 - \mathrm{tr}J + \mathrm{det}J = 0$. Since the flow is incompressible, $\partial_x u_x + \partial_y u_y = \mathrm{tr}J = 0$, thence $-\lambda^2 = \mathrm{det}J$. Given squared vorticity $\omega^2 = (\partial_x u_y - \partial_y u_x)^2$, and squared strain $s^2 = s_1^2 + s_2^2$, where the normal component is $s_1 = \partial_x u_x - \partial_y u_y$ and the shear component is $s_2 = \partial_y u_x + \partial_x u_y$, we may write

$$Q = \lambda^2 = -\mathrm{det}J = (s^2 - \omega^2)/4, \tag{21}$$

where Q is the Okubo–Weiss parameter ([17,24]). If $Q > 0$, $\lambda^2 > 0$, and λ is real, so deformation dominates, as around hyperbolic points, whereas if $Q < 0$, $\lambda^2 < 0$, and λ is complex, so rotation dominates, as near elliptic points. Equation (18) together with $d\mathbf{x}/dt = \mathbf{A} + \mathbf{u}$ defines a dissipative dynamical system

$$d\boldsymbol{\xi}/dt = \mathbf{F}(\boldsymbol{\xi}) \tag{22}$$

with constant divergence $\nabla\cdot\mathbf{F} = -4/3\,\mathrm{St}^{-1}$ in the four dimensional phase space $\boldsymbol{\xi} = (x, y, A_x, A_y)$, so that while small values of St allow for large values of the divergence, large values of St force the divergence to be small. The Stokes number is the relaxation time of the particle back onto the fluid trajectories compared to the time scale of the flow — with larger St, the particle has more independence from the fluid flow. From Eq. (20), about areas of the flow near to hyperbolic stagnation points with $Q > 4/9\,\mathrm{St}^{-2}$, particle and flow trajectories separate exponentially.

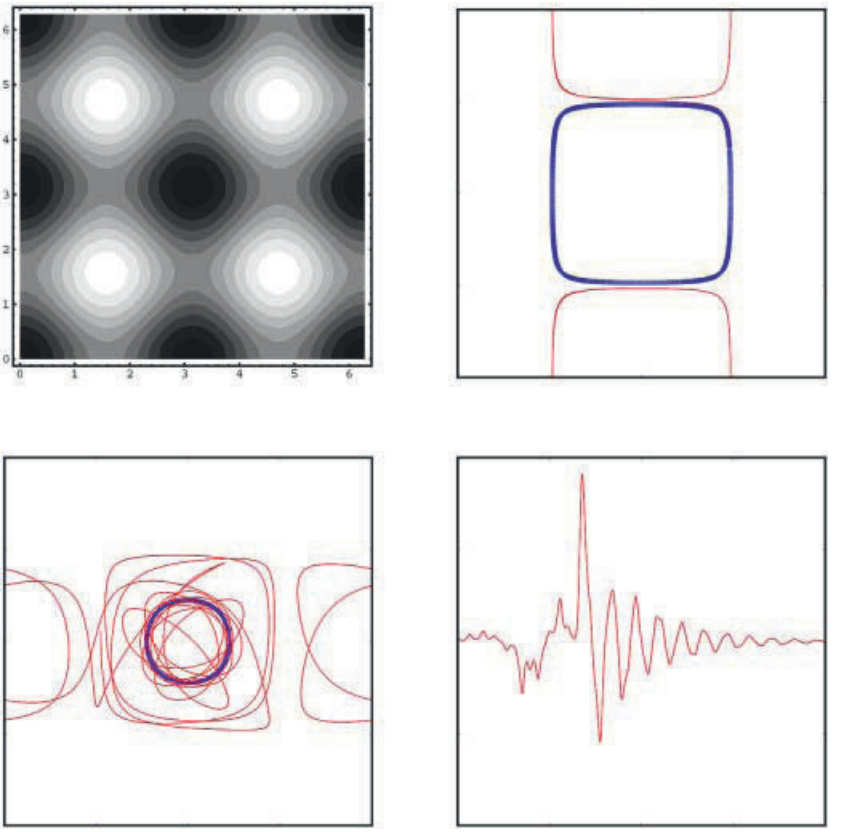

Fig. 2. (Top left) Contour plot illustrating magnitude of Q — lighter is higher Q — for the time-independent model Eq. (23) (the flow is on a torus). (Top right) The separation of a neutrally buoyant particle trajectory (thin line) from the flow (thick line) in regions of high Q allows the particle to wander between cells. (Bottom left) After a complicated excursion, a particle (thin line) eventually settles in a zone of low Q of the flow; a KAM torus (thick line). (Bottom right) The velocity difference $v_x - u_x$ between the particle and the flow against time.

To illustrate the effects of St and Q on the dynamics of a neutrally buoyant particle, let us consider the simple incompressible two-dimensional model flow defined by the stream function

$$\psi(x, y, t) = A \cos(x + B \sin \omega t) \cos y. \tag{23}$$

The equations of motion for an element of the fluid will then be $\dot{x} = \partial_y \psi$, $\dot{y} = -\partial_x \psi$. As we observed earlier, ψ has the rôle of a Hamiltonian for the dynamics of such an element, with x and y playing the parts of the conjugate coordinate and momentum pair. Let us first consider the simplest case where the time dependence is suppressed, by setting $B = 0$. Thence ψ should be a constant

of motion, which implies that real fluid elements follow trajectories that are level curves of ψ. In Fig. 2 are depicted contours of Q. Notice that the high values of Q are around the hyperbolic points, while negative Q coincides with the centres of vortices — elliptic points — in the flow. Figure 2 (top right) shows the trajectory of a neutrally buoyant particle starting from a point on a fluid trajectory within the central vortex, but with a small velocity mismatch with the flow. This mismatch is amplified in the vicinity of the hyperbolic stagnation points where Q is larger than $4/9\,\mathrm{S}t^{-2}$ to the extent that the particle leaves the central vortex for one of its neighbours, a trip that is not allowed to a fluid parcel. In the end a particle settles on a trajectory that does not visit regions of high Q, proper for a fluid parcel. While this effect is already seen at top right in Fig. 2b, it is more dramatically pictured in the trajectory shown at bottom left, in which the particle performs a long and complicated excursion wandering between different vortices before it settles in a region of low Q of one of them. To illustrate the divergence of particle and fluid trajectories, and the fact that particle and fluid finally arrive at an accord, at bottom right we display the difference between the particle velocity and the fluid velocity at the site of the particle against time for this case. Notice that this difference seems negligible at time zero, and that it also convergences to zero at long times, but during the interval in which the excursion takes place it fluctuates wildly.

Even more interesting is the case of time-dependent flow: $B \neq 0$ in our model. As in a typical Hamiltonian system, associated with the original hyperbolic stagnation points, there are regions of the phase space, which is here real space, dominated by chaotic trajectories. Trajectories of this kind, stroboscopically sampled at the frequency of the flow, are reproduced in Fig. 3. Such trajectories visit a large region of the space, which includes the original hyperbolic stagnation points and their vicinities where Q is large. Excluded from the reach of such a chaotic trajectory remain areas where the dynamics is regular; the so-called KAM tori. In our model these lie in the regions where $Q < 4/9\,\mathrm{S}t^{-2}$. Now a neutrally buoyant particle trying to follow a chaotic flow pathline would eventually reach the highly hyperbolic regions of the flow. This makes likely its separation and departure from such a pathline, in search of another pathline to which to converge. However, convergence will only be achieved if the pathline never crosses areas of high Q. Figure 3 demonstrates this phenomenon: a particle was released in the chaotic zone with a small velocity mismatch. The particle followed the flow, until, coming upon a region of sufficiently high Q, it was thrown out of that flow pathline onto a long excursion that finally ended up in a regular region of the flow on a KAM torus. The regular regions of the flow then constitute attractors of the dissipative dynamical system Eq. (22) that describes the behaviour of a neutrally buoyant particle. The chaotic trajectories in a Hamiltonian system are characterized by positive Lyapunov exponents. The Lyapunov exponents are an average along the trajectory of the local rate of convergence or divergence. Such a rate is measured by the quantity λ. Hence, for a trajectory to be chaotic, it is a necessary condition that it visit regions of positive Q: an upper bound to Q is an upper bound to the Lyapunov exponent.

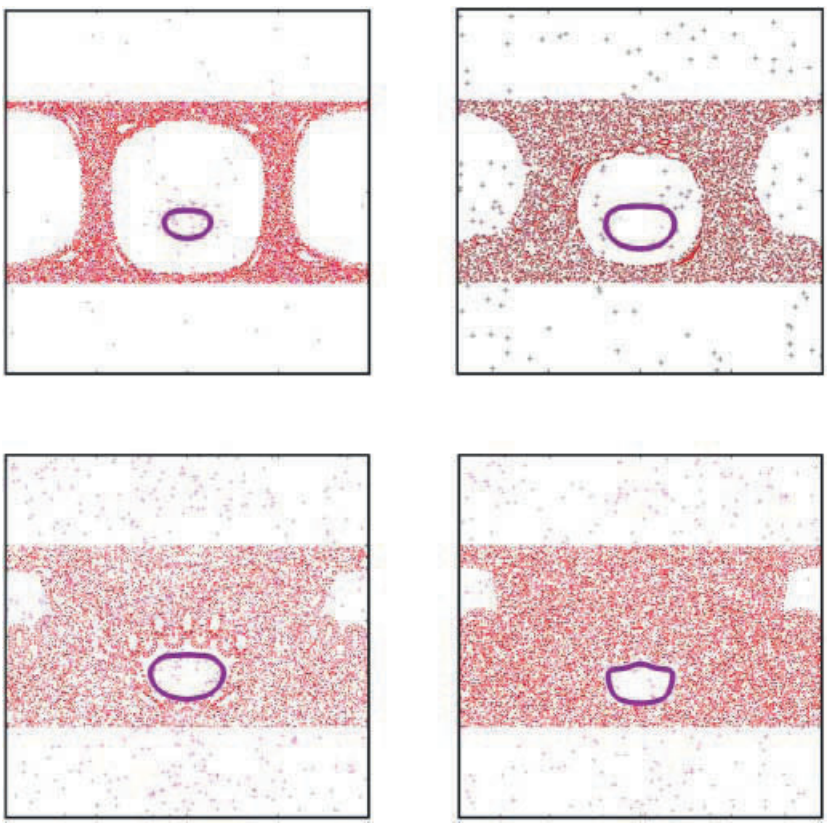

Fig. 3. Poincaré sections of trajectories in the time-dependent flow of Eq. (23). From top left to bottom right are shown (dots) four increasingly chaotic examples of the flow, and (crosses), the trajectories of neutrally buoyant particles in the flows that in each case finally end up on a KAM torus within the regular region of the flow.

We have considered the implications of these results for two-dimensional turbulent flows, in which Q defines three regions: in the vortex centres it is strongly negative; in the circulation cells that surround them, stongly positive, while in the background between vortices it fluctuates close to zero (see, e.g., [12,3,14,19,20]). As a result of the dynamics, an initially uniform distribution of neutrally buoyant particles with finite size evolves in time towards an asymptotic distribution concentrated in the inner part of vortices where $Q < 0$, and with voids in the areas crossed by fluid trajectories that visit regions where $Q > 4/9\,\mathrm{St}^{-2}$, as we illustrate in Fig. 4.

We have shown with the simplest model for the force acting on a small rigid neutrally buoyant spherical tracer particle in an incompressible two-dimensional fluid flow that the tracer trajectories can separate from the fluid trajectories in those regions where the flow has hyperbolic stagnation points. For flows with

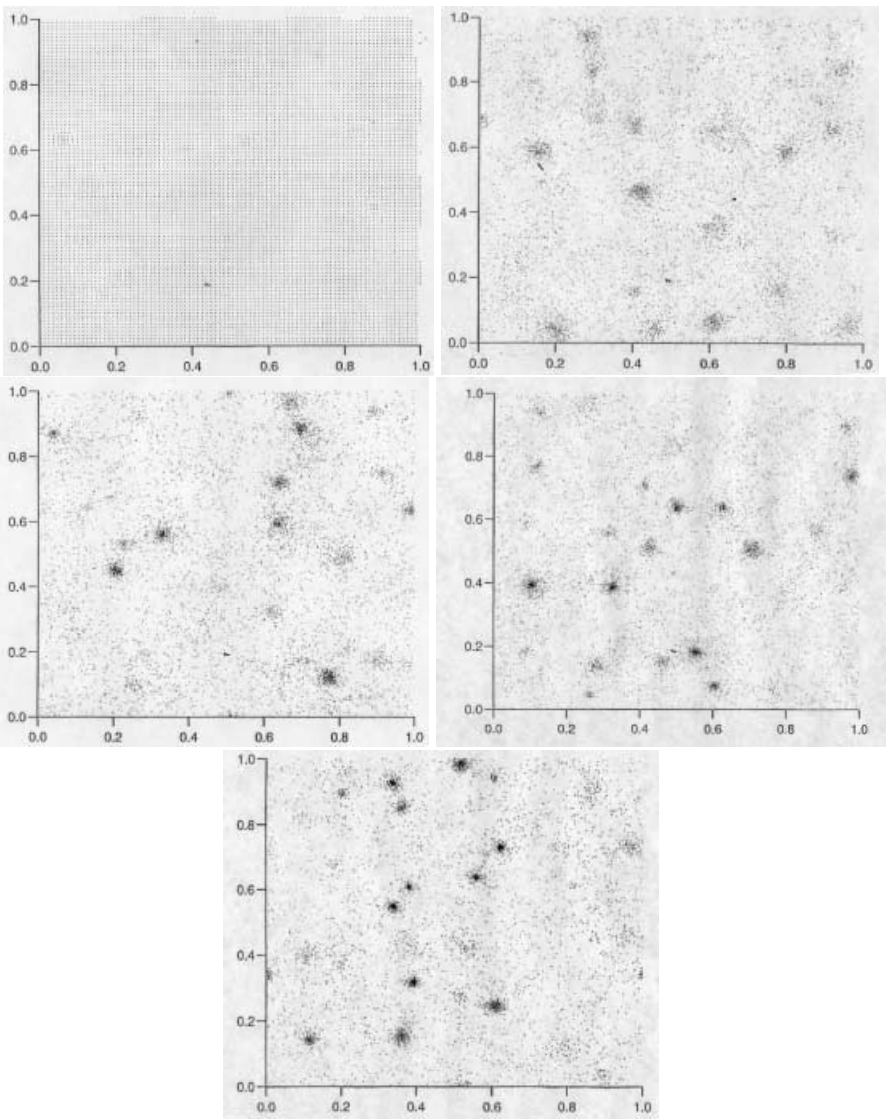

Fig. 4. Small neutrally buoyant tracer particles collect in the centres of vortices in a two-dimensional turbulent flow simulation; distribution at times (top left to bottom right) $t = 1$, 2, 3, 4, and 6 of particles uniformly distributed in the flow at time $t = 1$.

chaotic pathlines, analysis shows that the tracer will only evolve on trajectories having Lyapunov exponents bounded by the value of the Stokes drag coefficient. Therefore, by making the value of this coefficient small enough, one can force the tracer to settle on either the regular KAM-tori dominated regions or to selectively visit the chaotic regions with small Lyapunov exponents.

As well as its interest from the viewpoint of dynamical systems, this result should have some impact on the analysis of observations and experiments with neutrally buoyant particles both in the laboratory and in the atmosphere and oceans.

Acknowledgements

This review is based in part on our recent article [2].

References

1. Auton T. R., Hunt J. C. R., and Prud'homme M. (1988): The force exerted on a body in inviscid unsteady non-uniform rotational flow. J. Fluid Mech. **197**, 241–257
2. Babiano A. , Cartwright J. H. E., Piro O., and Provenzale A. (2000): Dynamics of a small neutrally buoyant sphere in a fluid and targeting in Hamiltonian systems. Phys. Rev. Lett. **84**, 5764–5767
3. Basdevant C. and Philipovich T. (1994): On the validity of the Weiss criterion in two-dimensional turbulence. Physica D **73**, 17–30
4. Basset A. B. (1888): On the motion of a sphere in a viscous liquid. Phil. Trans. Roy. Soc. **179**, 43–63
5. Boussinesq J. (1885): Sur la resistance qu'oppose un liquide indefini en repos. C. R. Acad. Sci. Paris **100**, 935–937
6. Cartwright J. H. E., Feingold M., and Piro O. (1994): Passive scalars and three-dimensional Liouvillian maps. Physica D **76**, 22–33
7. Cartwright J. H. E., Feingold M., and Piro O. (1995): Global diffusion in a realistic three-dimensional time-dependent nonturbulent fluid flow. Phys. Rev. Lett. **75**, 3669–3672
8. Cartwright J. H. E., Feingold M., and Piro O. (1996): Chaotic advection in three-dimensional unsteady incompressible laminar flow. J. Fluid Mech. **316**, 259–284
9. Cartwright J. H. E., Feingold M., and Piro O. (1999): An introduction to chaotic advection, in *Mixing: Chaos and Turbulence*, Eds. Chate H., Villermaux E., and Chomez J. M. (Kluwer) 307–342
10. Crisanti A., Falcioni M., Provenzale A., and Vulpiani A. (1990): Passive advection of particles denser than the surrounding fluid. Phys. Lett. A **150**, 79–84
11. Druzhinin O. A. and Ostrovsky L. A. (1994): The Influence of the Basset Force on Particle Dynamics in Two-Dimensional Flows. Physica D **76**, 34–43
12. Elhmaidi D., Provenzale A., and Babiano A. (1993): Elementary topology of two-dimensional turbulence from a Lagrangian viewpoint and single-particle dispersion. J. Fluid Mech. **257**, 533–558
13. Faxén H. (1922): Der Widerstand gegen die Bewegung einer Starren Kugel in einer zähen Flüssigkeit, die zwischen zwei parallen Ebenen Wänden eingeschlossen ist. Ann. Phys. **4**, 89–119
14. Hua B. L. and Klein P. (1998): An exact criterion for the stirring properties of nearly two-dimensional turbulence. Physica D **113**, 98–110
15. Maxey M. R. and Riley J. J. (1983): Equation of motion for a small rigid sphere in a nonuniform flow. Phys. Fluids **26**, 883–889

16. Michaelides E. E. (1997): Review — the transient equation of motion for particles, bubbles, and droplets. J. Fluids Eng. **119**, 233–247
17. Okubo A. (1970): Horizontal dispersion of floating particles in the vicinity of velocity singularities such as convergencies. Deep-Sea Res. **17**, 445–454
18. Ottino J. M. (1989): *The Kinematics of Mixing: Stretching, Chaos, and Transport* (Cambridge University Press)
19. Protas B., Babiano A., and Kevlahan N. K.-R. (1999): On geometrical alignment properties of two-dimensional forced turbulence. Physica D **128**, 169–179
20. Provenzale A. (1999): Transport by coherent barotropic vortices. Annu. Rev. Fluid Mech. **31**, 55–93
21. Russel W. B., Saville D. A., and Schowalter W. R. (1989): *Colloidal Dispersions* (Cambridge University Press)
22. Tanga P. and Provenzale A. (1994): Dynamics of advected tracers with varying buoyancy. Physica D **76**, 202–215
23. Taylor G. I. (1928): The forces on a body placed in a curved or converging stream of fluid. Proc. Roy. Soc. A **120**, 260–283
24. Weiss J. B. (1991): The dynamics of enstrophy transfer in two-dimensional hydrodynamics. Physica D **48**, 273–294
25. Yannacopoulos A. N., Rowlands G., and King G. P. (1997): Influence of particle inertia and Basset force on tracer dynamics: Analytical results in the small-inertia limit. Phys. Rev. E **55**, 4148–4157

Nonlinear Dynamics of Wrinkled Premixed Flames and Related Statistical Problems

Guy Joulin[1], Gaël Boury[1], Pierre Cambray[1], Yves D'Angelo[1], and Karl Joulain[2]

[1] Laboratoire de Combustion et de Détonique, UPR 9028 du CNRS, ENSMA, 86960 Futuroscope, Poitiers, France.
[2] Laboratoire EM2C, Ecole Centrale de Paris, 92290 Châtenay–Malabry, France.

Abstract. The flames propagating through premixed gaseous reactants are surface–like interfaces, that consume the former at a local curvature–dependent burning speed, and heat them. The Michelson–Sivashinsky (MS) equation and its variants well describe how the fluid–mechanical (Landau–Darrieus) instability competes with curvature and geometrical (Huygens) non–linearity in the overall flame shape dynamics. For nearly–straight or nearly–circular fronts, the dynamics translates into N–body problems for complex singularities. Approximate mean–field treatments yield predictions on the statistics of spontaneous wrinkling, that are checked against numerical integrations of MS equations : mean wrinkle–wavelength evolutions, etc... The paper next addresses how flames respond to weak broad–banded forcing, e.g. residual turbulence. Open questions relating to 3–D flames, or to the influence of a large tangential velocity, are finally evoked.

1 Introduction

We summarize below a few key concepts, and the associated notation, relating to the physics of premixed gaseous flames, for future use in the main body of the paper. It'll successively adress the nonlinear dynamics of flat-on-average fronts (Sections 2 to 4), the influence of noise on the latter (Sec. 5), then the specificities of expanding flames (Sec. 6). We end up with concluding remarks and a few open problems (Sec. 7).

1.1 Basics of Premixed Flames

Consider gaseous reactants in adequate proportions, e.g. city–gas and air, that are initially mixed homogeneously at the molecular level. Intuition suggests – and experiments confirm – that such mild ignition sources as matches or sparks will trigger a rapid conversion of the initial medium into burned, hotter gases. The process is not instantaneous ; instead, the conversion can be seen (by a camera, say) to take place across a thin combustion wave – the flame front – that propagates itself at a markedly subsonic burning speed relative to the medium just ahead of it. Let S_n and l denote typical values of that burning speed and of the front thickness, respectively. If t_{ch} represents a typical chemical time, e.g. based upon the maximum rate of heat release inside the front, and

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 127–158, 2001.
© Springer-Verlag Berlin Heidelberg 2001

D_{th} is the medium heat diffusivity, one may estimate l and S_n by $l = S_n t_{ch}$ (the mixture burns when crossing the front), and $l^2 = D_{th} t_{ch}$ (the heat released by chemistry is *conducted* into the fresh medium). Accordingly, $S_n \sim (D_{th}/t_{ch})^{1/2}$ and $l \sim D_{th}/S_n$. Reactive collisions between reactants being rare events, S_n is indeed *markedly subsonic* (S_n usually ranges from a few cm/s to some 10 m/s) and, since D_{th} is usually small, flames are fairly *thin* ($l = 10^{-6}$ to 10^{-4} meter, say). S_L will denote the value of S_n when the front is flat and propagates steadily in a homogeneous premixture at rest (in a suitable frame).

1.2 Curvature Effects

Flames seldom stay flat, for reasons explained later on. This has two main consequences. Firstly, the fresh gases acquire a nonuniform velocity field $\mathbf{u}$ in any one frame, which necessitates to define S_n accordingly, as

$$S_n \equiv \mathbf{n} \cdot (\mathbf{u} - \mathbf{D})_{fresh} \tag{1}$$

where $\mathbf{D}$ is the velocity (in the same frame) of the front, and the unit normal $\mathbf{n}$ to it points towards the burned gases. Implicit in (1) is the assumption that the thin front is idealized as a surface.

A second consequence of front wrinkling and of $\mathbf{u} \neq 0$ is that S_n no longer equals S_L. Theories [1] – [2] and measurements [3] – [4] agree on the law

$$\frac{S_n}{S_L} \;=\; 1 - \mathcal{L}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) + \frac{\mathcal{L}}{S_L}\mathbf{n} \cdot (\nabla \mathbf{u})_{fresh} \cdot \mathbf{n} \tag{2}$$

for most reactants, lean mixtures of H_2 being one of the few exceptions. In (2) the R_i's are the principal radii of curvature of the front surface (> 0 if the latter is convex towards the fresh side) and $\nabla \mathbf{u}$ is the rate–of–strain tensor just ahead of it. The Markstein length $\mathcal{L}$ [5], usually positive, is proportional to D_{th}/S_L, but often exceeds it by a significant numerical factor (≥ 5) ; furthermore, both contributions to $\dfrac{S_n}{S_L} - 1$ add up and have the same form for spontaneous evolutions thereby effectively doubling the net effect of curvature, approximately. Thus, one may simultaneously account for $\mathcal{L} \neq 0$ and consider the front as thin, owing to this $\gtrsim 10$ numerical factor.

1.3 Landau–Darrieus Instability

One of the main goals of studies on premixed–flame is to determine the front's normal displacement-speed $\mathbf{n}.\mathbf{D}$ in a given frame, for prescribed operating conditions. If $\mathbf{u}$ in (1) corresponded to a uniform flow, or actually were any prescribed field, combining (1) and (2) would yield an eikonal equation, possibly forced, for the front shape. Unfortunately $\mathbf{u}$ is *not* prescribable beforehand, as it depends on the flame shape evolution via the fluid mechanics in the fresh or burnt gases and their different *densities* (ρ_u or $\rho_b < \rho_u$, respectively).

For the sake of illustration one provisionally forgets about all finite–flame–thickness effects ($\mathcal{L}S_L/D_{th} = 0$), and hence all diffusive transports. The front is then a genuine surface equipped with a presumed known $S_n \equiv S_L$, and separating piecewise incompressible media. Next let us consider the simple case where the front is nearly planar, with an equation $z = F(t, \mathbf{x}) \ll 1$ in the frame of reference ($z \equiv 0$) of a steady flat front propagating along the z-axis ; here and below t denotes time and $\mathbf{x}$ are transverse Cartesian coordinates. The approximation $|\nabla F| \to 0$ enables one to linearize the Euler equations (about $\overline{\mathbf{u}} = (S_L, 0)$ or $(S_L\rho_u/\rho_b, 0)$), the Hugoniot relationships [6] at the front, then the kinematic relationship (1). Next one may restrict $F(t, \mathbf{x})$ to $\mathcal{F}\exp(\omega t + i\mathbf{k}\cdot\mathbf{x})$, where $\mathcal{F}|k| \ll 1$, the transverse wavevector $\mathbf{k}$ is given and the growth/decay rate ω has to be determined. Dimensional analysis and isotropy considerations imply [7]

$$\omega = |k| S_L \Omega(\gamma), \qquad |k| = (\mathbf{k}\cdot\mathbf{k})^{1/2}, \qquad \gamma \equiv \frac{\rho_u - \rho_b}{\rho_u} \tag{3}$$

because the problem has no fixed reference length (only a speed, S_L) and hence is essentially *scale–invariant*. Explicit resolution [7] [8] yields $\Omega(\gamma)$ as the positive root of

$$(2-\gamma)\Omega^2 + 2\Omega = \frac{\gamma}{1-\gamma} \tag{4}$$

In such a crude model, all flames are unstable because γ is always positive (and < 1). As (1) specializes here to

$$\frac{\partial F}{\partial t} = (\mathbf{n}\cdot\mathbf{u})_{z\,=\,0^-} - S_L \left(\equiv u'(t, \mathbf{x}, 0^-)\right) \tag{5}$$

the instability is of purely fluid–mechanical origin. It is such that $u' > 0$ (resp. < 0) wherever $\mathcal{F} > 0$ (resp. < 0), thus leading to wrinkle amplification, and results from streamline deflection in the incompressible ($\rho \equiv \rho_u$) fresh medium. In absence of any large–scale stabilizing mechanisms (walls, buoyancy effects, ...) this Landau–Darrieus (LD) instability is *unavoidable whenever* $|k|\mathcal{L}$ *is small enough to yield* $S_n \simeq S_L$. A final remark : once (3) is used, (5) may also be written as

$$\frac{\partial F}{\partial t} = S_L \Omega(\gamma) I(F, \mathbf{x}) \tag{6}$$

where the linear operator $I(F, \mathbf{x})$ is the multiplication by $|k|$ in $\mathbf{x}$–wise Fourier space.

1.4 Dispersion Relation, $K_{neutral}$

Naively retaining $S_L\Omega(\gamma)I(F, \mathbf{x})$ as an estimate for $u'(t, \mathbf{x}, 0^-)$, next computing the corresponding $\mathbf{n}.\nabla\overline{\mathbf{u}}_{fresh}.\mathbf{n}$ ($=-S_L\Omega(\gamma)\nabla^2 F$) , and *now* retaining (2) in the linearized form of (1) yields

$$\omega = S_L \Omega |k| \left(1 - \frac{|k|}{K_{neutral}}\right) \tag{7}$$

with $\mathcal{L}K_{neutral} \equiv \dfrac{\Omega}{1+\Omega}$ here. Crude as it is, this reasoning qualitatively explains what brings about a *cut–off wavenumber* $K_{neutral}$ (see Fig. 1) in the dispersion relation $\omega(\mathbf{k})$: wherever $\mathcal{F} > 0$, i.e $u'(t, \mathbf{x}, 0^-) > 0$ (resp. < 0), S_n increases (resp. decreases), counteracting the Landau–Darrieus mechanism or even killing it if $|k| > K_{neutral}$. The above estimate also yields $lK_{neutral} \lesssim 0.1$ if $\mathcal{L}/l \geqslant 5$, as Ω usually is about unity ; $\Lambda_{neutral} \equiv 2\pi/K_{neutral}$ then exceeds $60l$.

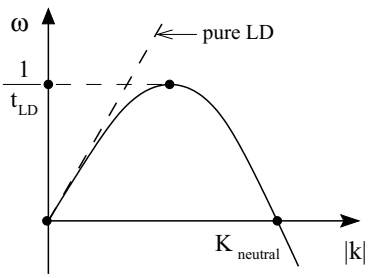

Fig. 1. Dispersion relation $\omega(\mathbf{k})$ for nearly–planar, wrinkled fronts, Eq. (7)

However, the procedure leading to (7) was not consistent, even though ω should always have the form $S_L|k|\Omega(\gamma, |k|l)$ on dimensional grounds. Systematic procedures that exploit the smallness of $|k|l$ over the range of unstable wavenumbers $|k|$ have been worked out [9] – [11], retaining realistic transports inside the flame front, viscosity effects outside, and the modifications of the Hugoniot relationships that are needed to be consistent with (2). The resulting ωs are very close to (7), yet with a different (and analytically determined) $K_{neutral}$. For realistic values of $\gamma = (\rho_u - \rho_b)/\rho_u$, $\gamma = 0.8$–0.9 say, $lK_{neutral}$ ranges from ~ 0.1 (lean CH_4/air) to ~ 0.05 (lean flames of heavy fuels, e.g. C_3H_8) [11].

1.5 Huygens Type of Nonlinearity

Only in the linear approximation does $\mathbf{n} \cdot \mathbf{D}$ coincide with $\partial F/\partial t$, in the same frame as above. Actually,

$$\mathbf{n} \cdot \mathbf{D} \equiv \frac{\partial F/\partial t}{\sqrt{1 + |\nabla F|^2}} \tag{8}$$

Following Ref [12] consider a last crude model where all fluid-mechanical effects are ignored. Then, for small $|\nabla F|^2$, (1) and (8) would yield

$$\frac{\partial F}{\partial t} \simeq -\frac{S_L}{2}|\nabla F|^2 + S_L \mathcal{L} \nabla^2 F \tag{9}$$

i.e. a Burgers equation for $F(t, \mathbf{x})$. This somehow suggests that *geometrical nonlinearities* will likely play an important role in the dynamics of premixed flames.

2 Michelson–Sivashinsky Equation(s)

2.1 Heuristics

Comparing (6) (7) (9) led Sivashinsky [12] to surmise that what misses in (6) shows up in (9), and conversely. A model equation synthesizing the two would then look like

$$\frac{\partial F}{\partial t} \simeq -\frac{S_L}{2}|\nabla F|^2 + S_L \mathcal{L}\nabla^2 F + S_L \Omega(\gamma) I(F,\mathbf{x}) \tag{10}$$

i.e. a mere "sum" of a Burgers equation and of the Landau–Darrieus result !

2.2 Small–γ Expansions

An equation resembling (10) is likely to emerge from a rational procedure if the three terms in the RHS are small and comparable, so that any cross–terms would be even smaller. This requires $\Omega(\gamma)$, hence γ, to be small, as
$2\Omega(\gamma \ll 1) = \gamma(\gamma/2 + ...)$, then

$$\frac{F^2}{x^2} \sim \frac{\mathcal{L}F}{x^2} \sim \frac{\gamma F}{x} \sim \frac{F}{S_L t} \tag{11}$$

A consistent balance thus requires $x \sim \dfrac{\mathcal{L}}{\gamma} \gg l$, $t \sim \dfrac{\mathcal{L}}{\gamma^2 S_L} \gg \dfrac{l}{S_L}$, which corresponds to a *quasi–planar, quasi–steady* wrinkled front.
Next, as the z-wise and the x-wise ranges of hydrodynamics are comparable and large compared to $F(t,\mathbf{x})$, matched–asymptotic expansions [12] allow one to handle the problem. Finally, as $u'(t,\mathbf{x},0^{\pm}) \sim \gamma^2 S_L \ll S_L$ and $x \gg l$, the hydrodynamical zones that flank the thinner front are governed by linearized Euler equations, to leading order in γ.

To first order in $\gamma \longrightarrow 0$, the equation for $F(t,\mathbf{x})$ is found [12] to be :

$$\frac{\partial F}{\partial t} + \frac{S_L}{2}|\nabla F|^2 = S_L \frac{\gamma}{2}\left(\frac{\nabla^2 F}{K_{neutral}} + I(F,\mathbf{x})\right) \tag{12}$$

with $K_{neutral}\mathcal{L} = \gamma/2$ here, i.e. precisely what (10) suggested. Equations with the same structure as (12) are known as **Michelson–Sivashinsky** (MS) equations.

Conducting the analysis to second order in γ led [13] [14] to

$$\frac{\partial F}{\partial t} + \frac{S_L}{2}|\nabla F|^2 a(\gamma) + \frac{S_L}{2}\left(1 - a(\gamma)\right)\langle|\nabla F|^2\rangle = S_L \frac{\gamma}{2}\left(1+\frac{\gamma}{2}\right)\left(\frac{\nabla^2 F}{K_{neutral}} + I(F,\mathbf{x})\right) \tag{13}$$

with a slightly different $K_{neutral}$ and $a(\gamma) = 1 + \gamma/2 > 1$.

Apart from modified coefficients that do not alter its structure, the main difference with (12) is the "counter–term" proportional to the transversal average $\langle|\nabla F|^2\rangle$; its presence is needed to ensure that the mean fractional increase in effective burning speed due to wrinkling $-\langle\partial F/\partial t\rangle/S_L$ coincides to the mean

increase in front area, viz. : $\langle(1+|\nabla F|^2)^{1/2}-1\rangle = \langle|\nabla F|^2\rangle/2+...$ Even though such counter–terms can be formally removed from the evolution equations upon properly redefining F, their need practically ruins the amplitude expansions proposed in [17] [18] as alternatives to the small–γ asymptotics, at least in the form published to date.

The third–order evolution equation [15] differs from (13) by several points : (i) $K_{neutral}$ is slightly different, as is the expression of the required counter–term. (ii) $a(\gamma)$ is modified into $1+\gamma/2+3\gamma^2/4$, whereas the coefficient in front of the RHS of (13) now reproduces $\Omega(\gamma)$ to third order in γ.
(iii) More importantly, the RHS of the new equation also includes 8 (sic) extra nonlinear terms,stemming from the expansion of $1/R_1+1/R_2$ in (2), from a more accurate description of the hydrodynamical fields, and from geometry.

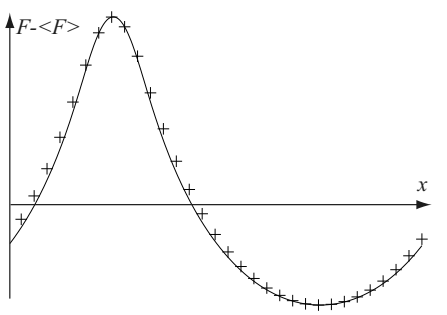

Fig. 2. Spatially periodic ($K = 0.5K_{neutral}$) flame front shapes for $\gamma = 0.8$. *Solid line* : computed from the 3^{rd}–order evolution equation for $F(t,x)$. *Symbols* : from the MS equation (14), with $a(\gamma)$ fitted to yield the same crest–to–trough amplitude [15]

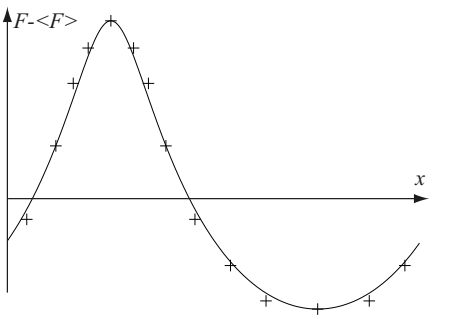

Fig. 3. Comparison between the steady solution of (14) (*Solid line*) and the results (*Symbols*) of a direct numerical simulation [20]. Both have $\gamma = 0.8$ and $K = 0.5K_{neutral}$. The $a(\gamma)$ used in Eq. (14) is the same as in Fig. 2

It was thus very surprising that the solutions to such a complicated equation turn out to be very close [15] (see Fig. 2) to those of

$$\frac{\partial F}{\partial t}+\frac{S_L}{2}|\nabla F|^2 a(\gamma)+\frac{S_L}{2}\left(1-a(\gamma)\right)\langle|\nabla F|^2\rangle = S_L\Omega(\gamma)\left(\frac{\nabla^2 F}{K_{neutral}}+I(F,\mathbf{x})\right) \tag{14}$$

even for *non–small* γ, provided $a(\gamma)$ is adequately chosen **once for all** as a function of γ and Ω is computed as the (unexpanded–) positive root of (4) ! The agreement holds for both steady and unsteady flames (see Fig. 3) and seems to extend [15] [16] to comparisons with the direct numerical simulations of [19] [20] and [16], for reasons yet to be understood ! That (14) fairly well reproduces the linear results is possibly not alien to the coincidence. Accordingly, the present essay will henceforth focus on (14) and simple variants thereof.

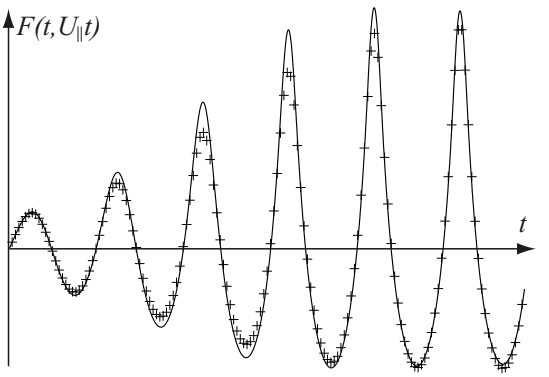

Fig. 4. Transient flame shapes $F(t,x)_{x=U_{\|}t}$ corresponding to $U_{\|} > \Omega(\gamma)S_L$, $\gamma = 0.8$, $K = 0.5K_{neutral}$. *Solid line* : from the 3^{rd}–order evolution equation. *Symbols* : from Eq.(14), with the same $a(\gamma)$ as in Figs. 2-3-3

3 Pole Dynamics

3.1 A Miracle

Equation (14) has the further advantage that infinitely many exact solutions to it can be obtained analytically when only one transverse coordinate (x) is involved. As noticed in [21], (14) indeed belongs to a class [22] of model equations that admits

$$F(t,x) = h(t) - \frac{2\Omega}{aK_{neutral}} \sum_{\alpha=1}^{2N} \mathrm{Log}\,\sin\left(\frac{K(x - x_\alpha(t))}{2}\right) \tag{15}$$

as exact $2\pi/K$–periodic ($K > 0$) solutions whatever Ω, a, K and integer $N \geq 0$. The complex numbers $x_1(t)$, $x_2(t)$, ..., $x_{2N}(t)$ appear in conjugate pairs in (15) for $F(t,x)$ to be real when x is, and can be identified with the poles of $\partial F/\partial x$ in the complex x-plane. For (15) to exactly solve (14) the x_α's must satisfy the $2N$ coupled complex ODEs :

$$\frac{dx_\alpha}{dt} = \Omega S_L \left[\frac{K}{K_{neutral}} \sum_{\beta\neq\alpha}^{2N} \mathrm{cotan}\left(\frac{K(x_\beta - x_\alpha)}{2}\right) - i\varepsilon_\alpha\right] \tag{16}$$

where $\varepsilon_\alpha = \mathrm{sign}\,(\mathcal{I}(x_\alpha))$, $\mathcal{I}(.)$ denoting an imaginary part ; furthermore, $\dfrac{dh}{dt} = -2\dfrac{\Omega^2}{a^2} S_L \dfrac{NK}{K_{neutral}} \left(1 - \dfrac{NK}{K_{neutral}}\right)$ in (15). That all poles are simple and have the same residue ($= -2\Omega K/aK_{neutral}$), hence are *undistinguishable*, results from identical local ($x \simeq x_\alpha$) dominant balances between nonlinearity and dissipation, like in Burgers' equation [23]. This undistinguishability has the consequence that it is very difficult to identify the physical origin of a front wrinkle one sees at time t along the real x-axis ! The imaginary convection ($\sim -i\varepsilon_\alpha$)

tends to push the poles towards the real axis, and comes from the $I(F,x)$ operator, whereas the interaction term in (16) (a remnant of $(\partial F/\partial x)^2$) counteracts this trend and prevents real singularities from appearing along the front. To ease physical interpretations of the results obtained from pole–dynamics, one may quote that each pair $A \pm iB$ of complex–conjugate poles yields the additive contribution

$$F_{AB}(t,x) = -\frac{2\Omega}{aK_{neutral}} \mathrm{Log}\left(1 - \frac{\cos K(x-A)}{\cosh KB}\right) \tag{17}$$

to the front shape, up to a shift. In the limit of infinitely long x–wise periods, where the $\sin(w)$ and $\mathrm{cotan}(w)$ functions are respectively replaced by w and $1/w$, (17) corresponds to what will be called an "elementary crest" Fig. 4. The

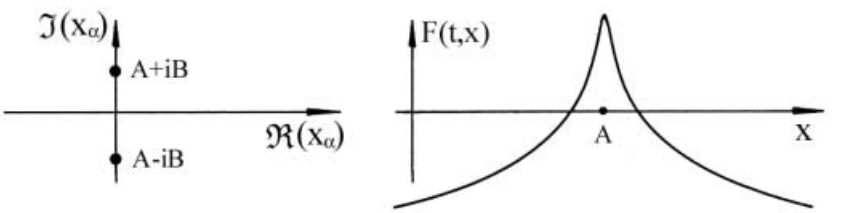

Fig. 5. *Left* : pair of poles $A \pm iB$ in complex x–plane ; *Right* : associated elementary crest pointing towards the burned gases

pole–dynamics (16) specifies the collective behavior of such interacting "solitons" (recall that N, when finite, is *conserved* in (15) (16)) and one of the goals of the analyses to be summarized next (subsection 4.3) is to get insights into it, even if crude ones.

3.2 More on Pole–Dynamics

As first shown in [21] specializing (16) to the case where two poles, x_1 and x_2, are close to each other in the same half complex x plane yields

$$\frac{d}{dt}(x_1 - x_2) \simeq \frac{4\Omega S_L}{K_{neutral}} \frac{1}{x_2 - x_1} \tag{18}$$

whereby x_1 and x_2 travel along hyperbolæ in their "center of mass frame", being subjected to mutual attraction (resp. repulsion) along the real (resp. imaginary) axis. The complex conjugates simultaneously do the same in the other half plane. When looked at from the real axis this process of pole alignment along the imaginary axis is seen as the *coalescense of elementary crests* (Fig. 5), a combustion analogue of Plato's cave myth ! As time elapses this process leads to a single giant crest in each (real–) interval of $2\pi/K$–length. Using (16), then the fact that $\mathrm{cotanh}(u)$ for any $u > 0$, one can show [14] [21] that the maximum number

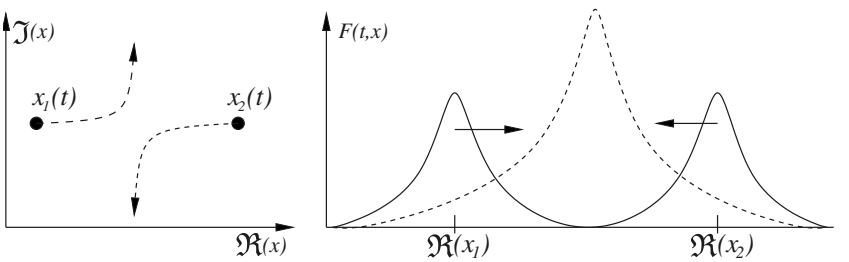

Fig. 6. *Left* : close encounter of two nearby poles $(x_1(t), x_2(t))$ in upper half complex x–plane ; *Right* : resulting crest coalescence as seen along the real x–axis

$N = N_{max}$ of *steady* aligned pairs of poles "belonging" to such a fully–coalesced steady giant crest is given by

$$N_{max} = \text{Int}\left(\frac{1}{2}(1 + \frac{K_{neutral}}{K})\right) \tag{19}$$

If $N > N_{max}$ initially, $N - N_{max}$ pairs of poles are ultimately rejected to $\pm i\infty$ and no longer contribute to the flame shape for $t \longrightarrow +\infty$. The fully–coalesced crests with $N = N_{max}$ have recently been shown [24] to represent the only stable (up to shifts) pole–decomposed steady solutions to (14). Two final facts about (19) are worth noticing :
(i) $N_{max} = 1$ for $1/3 < K/K_{neutral} \leq 1$ and, in particular, mature quasi–steady cells corresponding to the mostly amplified linear–modes ($K \simeq K_{neutral}/2$) have two poles "attached" to them ;
(ii) $N = N_{max}$ when the latter coalesce into larger cells, the mean pole number per unit real length is conserved to be $2/2\Lambda_{neutral}$ with $\Lambda_{neutral} \equiv 2\pi/K_{neutral}$, as (15) conserves N and coalescence is a mere rearrangement of the pole locations.

4 On the Statistics of Wrinkling

4.1 Mean Cell–Wavelengths from Numerics

Generating small broad–enough–banded initial conditions $F(0, x)$ from a suitable generator of pseudo–random numbers, and selecting large–enough domain sizes ($K \ll K_{neutral}$) one can integrate (14) numerically by a pseudo–spectral method (e.g. see [21] [25]). Repeating this for different generator seeds produces graphs like in figures 7 for the ensemble–averaged distance $\overline{\Lambda}(t)$ between the successive, usually sharp, maxima of $F(t, x)$ (see Figs. 8–9); each of them corresponds to a front–crest pointing towards the burnt gases.

In a first, linear, stage $\overline{\Lambda}(t)$ levels to about $2\Lambda_{neutral}$ as predicted from (7) because this corresponds to the mostly amplified mode of a nearly–planar front.

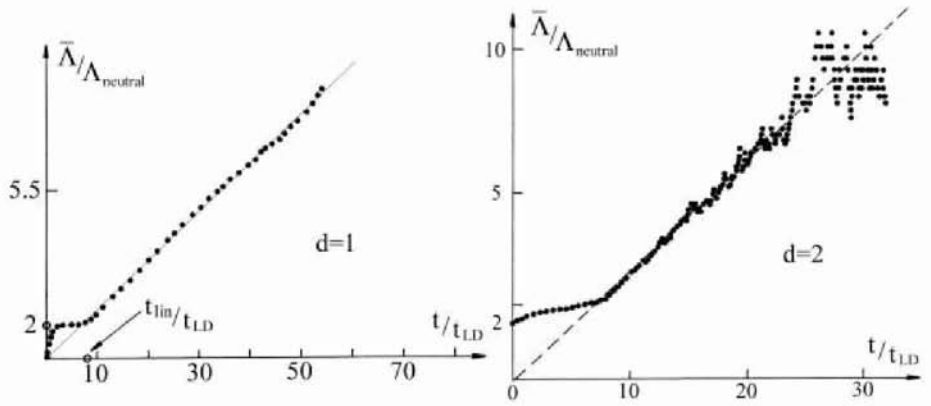

Fig. 7. Mean ensemble-averaged cell wavelength $\overline{\Lambda}(t)$ obtained numerically from (14), for one–dimensional (d=1) or two–dimensional (d=2) flame fronts. Periodic boundary conditions in both cases

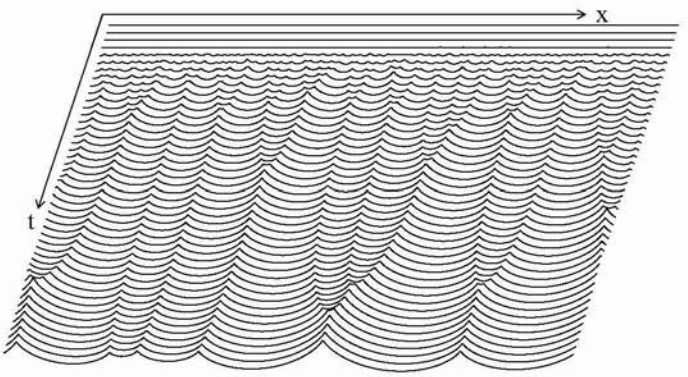

Fig. 8. Flame–shape evolution from a small random initial condition at t=0

The typical duration t_{lin} of the linear stage depends on the initial RMS deformation $\left(\overline{F^2(0,x)}\right)^{1/2}$, for its end corresponds to nonlinearity becoming significant, viz $\left(\overline{F^2(t_{lin},x)}\right)^{1/2} \sim \Omega/aK_{neutral}$; the cell wavelengths are then almost equidistributed, with a typical fractional wavelength dispersion of about $(t_{LD}/t_{lin})^{1/2}$, where

$$t_{LD} \equiv \frac{4}{\Omega S_L K_{neutral}} \tag{20}$$

is the minimum growth time associated with $\max(\omega)$ in (7).

Beyond $t \sim t_{lin}$, $\overline{\Lambda}(t)$ was found to grow *linearly* with time, the asymptotic slope $d\overline{\Lambda}(t)/dt$ being close to $\Omega S_L/2$ [25]. What happens in the limit of very long times is to be evoked in Sec. 5 below.

4.2 Heuristics

How to explain the above linear growth of $\overline{\Lambda}(t)$? Guided by the numerics (e.g. see Figs. 7), one might argue that the typical ratio $\overline{A}(t)/\overline{\Lambda}(t)$ between wrinkle amplitudes $\overline{A}(t)$ and cell–wavelengths $\sim \overline{\Lambda}(t)$ is nearly constant in the nonlinear regime where $\overline{\Lambda}(t) \sim t$ holds. This actually corresponds to a balance between $a(\partial F/\partial x)^2$ and $\Omega I(F,x)$ in (14). If true, this implies

$$\frac{d\,\mathrm{Log}(\overline{\Lambda})}{dt} = \frac{d\,\mathrm{Log}(\overline{A})}{dt} \tag{21}$$

Besides, on the scale $\overline{\Lambda} \gg \Lambda_{neutral}$ the not–yet–used terms in (14) yield $dA/dt \sim AS_L\Omega/\Lambda$, as the only scale left over for $\Lambda \gg \Lambda_{neutral}$ is $\overline{\Lambda}$ itself. Hence one "deduces"

$$\frac{d\overline{\Lambda}}{dt} \sim \Omega S_L \tag{22}$$

Another, seemingly different, line of thoughts is as follows. Viewing the process of crest coalescence as 2nd order chemistry : the concentration in crests $1/\overline{\Lambda}$ would decay at a rate proportional to $(1/\overline{\Lambda})^2$, the "rate constant" being itself proportional to $S_L\Omega$ for dimensional reasons and because it vanishes if Ω does. Again, this "results" in $\overline{\Lambda} \sim \Omega(\gamma)S_L t$.

All this suggests that $\overline{\Lambda} \sim S_L t$ is a consequence of the Landau–Darrieus instability **and** the dominant nonlinearity being **scale–invariant**. Then, how to theoretically estimate the asymptotic value of $\zeta \equiv (\Omega S_L)^{-1} d\overline{\Lambda}/dt$, a **pure** number apparently?

4.3 An Approximate, Mean–Field Recursion

The number ζ defined above lies beyond mere dimensional arguments (yet it was defined in such a context) and estimating its value (if it exists !) requires a deeper dive into the "mechanics" of crest interactions (see Fig. 11). Even in the context of pole–decomposed solutions to (14), nobody could so far solve the problem rigorously, the $2N$–body problem (16) being to date too difficult. A rough description [25] of the crest statistics is summarized below.

Repeated numerical investigations of (14) convinced us that :
(i) most of the time during the coalescence process, the population of singularities consists of "bunches" of P_m poles having nearly identical real parts $X_m(t)$ and comparatively smaller imaginary parts than the real distance to the nearest next bunch, (see Fig. 10);
(ii) when two "bunches" collide, they quickly form another one, the number of poles involved being conserved in the collision. This led to investigate the

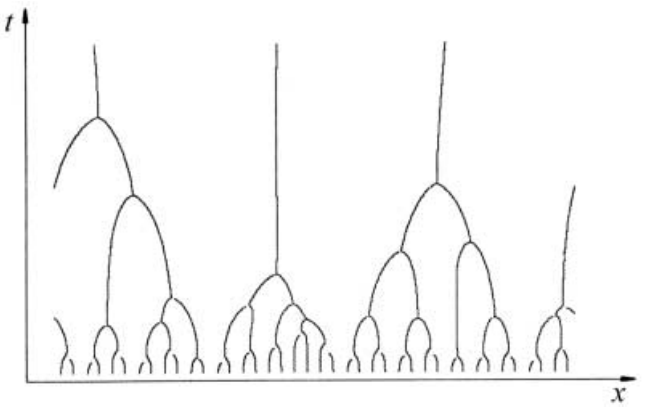

Fig. 9. Typical (x,t) crest trajectories showing the process of cell coalescence, in conditions similar to Fig. 8

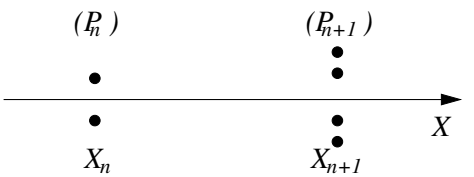

Fig. 10. Two "bunches" of poles, each associated with a front crest (located at $X = X_n, X_{n+1}$)

following restricted many–body problem along the whole real line :

$$\frac{dX_n}{dt} = \frac{2\Omega S_L}{K_{neutral}} \sum_{m \neq n} \frac{P_m}{X_m - X_n} \tag{23}$$

where $X_m(t)$ is the (real–)location of the m^{th}–crest and P_m is its "weight", i.e. the number of poles attached to it. The dynamics (23), obtained upon setting $K \to 0$ in (16), taking the real part thereof and neglecting the poles'imaginary parts, is to be supplemented by the Collision Rule (CR) :

$$\text{CR} = \text{"the weights of colliding crests add up at the collision"} \tag{24}$$

as to express that the poles are indestructible. Though simpler that (16) the inelastic [1] dynamics (23) (24) was still too difficult and further approximations were invoked. Focusing on two anonymous neighboring crests, located at

[1] Actually, equations (23) (24) may be lumped into $\frac{dX_n}{dt} \sim \lim_{\varepsilon \to 0^+} \sum_{m \neq n} f(X_m - X_n)$; $f(X) = X/(\varepsilon + X^2)$. Accordingly a Lyapunov function exists, which might help

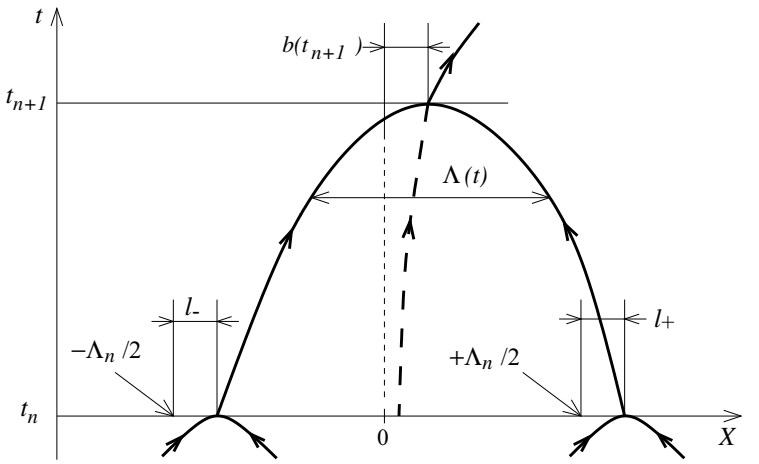

Fig. 11. A typical binary collision between crests showing the origin of a recursion between probability densities

$X_\pm(t_n) = \pm\overline{\Lambda}(t_n)/2 + l_\pm$ when $t = t_n$ (Fig. 11) and having weights $P_\pm$, we simplified their dynamics upon assuming that, during the approach to collision, all the other crests can be replaced by a mean continuous distribution with weight density $\dfrac{\overline{P}(t_n)}{\overline{\Lambda}(t_n)} = \dfrac{2}{2\Lambda_{neutral}}$, spreading from $|X| = 3\overline{\Lambda}(t_n)/2$ to $|X| \longrightarrow +\infty$. This choice of $\dfrac{\overline{P}(t_n)}{\overline{\Lambda}(t_n)}$ was guided by the remark ending subsection 3.2, and by the fact that the beginning of the coalescence phase corresponds to nearly equidistributed matures crests with $\Lambda \simeq 2\Lambda_{neutral}$ (hence $P = 2$ for each). Approximating next $\dfrac{P_+ + P_-}{2}$ by the mean crest weight $\overline{P}(t)$ evaluated at $t = t_n$, the wavelength $\Lambda(t) = X_+(t) - X_-(t)$ and the barycenter $b(t) = \dfrac{X_+(t) + X_-(t)}{2}$ were found to satisfy

$$\frac{d\Lambda}{dt} = \frac{4\Omega S_L \overline{P}(t_n)}{K_{neutral}} \left(-\frac{1}{\Lambda} + \frac{\Lambda}{\overline{\Lambda}(t_n)^2} \right) \tag{25}$$

with $\Lambda(t_n) = \overline{\Lambda}(t_n) + l_+ - l_-$, and

$$\frac{db}{dt} = \frac{4\Omega S_L \overline{P}(t_n)}{K_{neutral}} \frac{b}{\overline{\Lambda}(t_n)^2} \qquad 2b(t_n) = l_+ + l_- \tag{26}$$

Integration gave the time of collision $t = t_{n+1}$ and $B \equiv b(t_{n+1})$ as

$$t_{n+1} - t_n = \frac{K_{neutral}\overline{\Lambda}(t_n)^2}{8\Omega S_L \overline{P}(t_n)} \mathrm{Log}\left(\frac{\overline{\Lambda}(t_n)}{2(l_- - l_+)} \right), \text{ iff } l_- > l_+ \tag{27}$$

$$B = \frac{l_+ + l_-}{2} . \left(\frac{\overline{\Lambda}(t_n)}{2(l_- - l_+)} \right)^{1/2}, \text{ iff } l_- > l_+ \tag{28}$$

Considering $l_\pm$ as *statistically independent random* variables distributed according to the *same* even Pdf, $f_n(l_\pm)$, one can compute the probability of getting the collision location in $(B, B+dB)$. Acknowledging that B plays the same part at t_{n+1} as $l_\pm$ did at t_n, and that $\overline{\Lambda}(t_{n+1}) = 2\overline{\Lambda}(t_n)$, we got the functional recursion

$$F_{n+1}(z) = \int_{-\infty}^{+\infty} F_n\left(\frac{uz+u^2}{2}\right).F_n\left(\frac{uz-u^2}{2}\right)u^2 du \tag{29}$$

where $f_n(l)$ was written as $32F_n(z)/\overline{\Lambda}(t_n)$, and $z \equiv 32l/\overline{\Lambda}(t_n)$. This recursion has $\delta(z)$ as unstable fixed point corresponding to equidistributed crests (no coalescence will occur then) ; it also has an attractive one satisfying $F(+z) = F(-z)$, $F(|z| \gg 1) \sim 1/|z|^3$ and having an $\mathcal{O}(1)$ half–width... in units of $\overline{\Lambda}(t)/32$ (Fig. 12). From the successive F_n's and from (27) the mean collision time

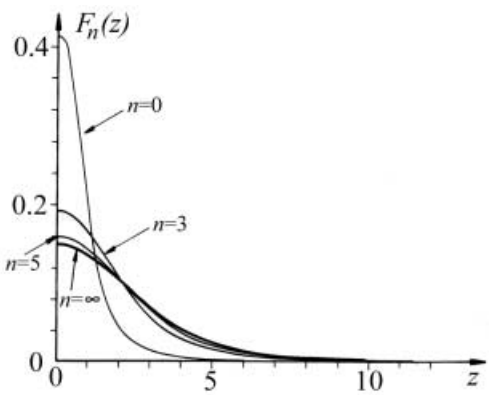

Fig. 12. Relaxation of the solutions to (29) towards the stable fixed point

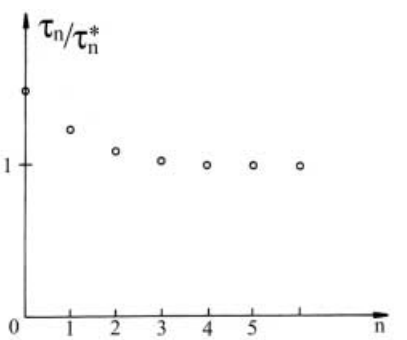

Fig. 13. Relaxation of τ_n/τ_n^* (Eq. (30)) towards unity

$\tau_n \equiv \overline{t_{n+1} - t_n}$ was found to quickly converge (Fig. 13) to a known value

$$\tau_n \longrightarrow \frac{\overline{\Lambda}(t_n)^2 K_{neutral}}{4\Omega S_L \overline{P}(t_n)} \equiv \tau_n^* \tag{30}$$

from *above* if one starts from a *narrow* $f_0(l_\pm)$.

From this it followed that $\dfrac{d\overline{\Lambda}}{dt} \simeq \dfrac{\overline{\Lambda}(t_{n+1}) - \overline{\Lambda}(t_n)}{\tau_n}$ approached $\zeta S_L \Omega$ from *below*, with a coefficient ζ exceeding that found in 4.1 by a factor of 1.3 only. One of the reasons why $d\overline{\Lambda}/dt$ is overestimated follows from having neglected the pole imaginary parts, which tends to shorten the collisions between crests. Computing the exact ζ is one of the many unsolved statistical problems appearing in nonlinear flame dynamics, especially when $F(0^+, x)$ does not belong to the class of functions (15), if it can [26] [27].

4.4 Spectra, Fractalization

Though an interesting exercise that led to a fair estimate of ζ, the approach above is too crude to give access to fine statistics of wrinkling, particularly for very long times. A more complete description would introduce the one–time, two–point ($\mathbf{x}$ and $\mathbf{x}'$) ensemble–autocorrelation $\mathcal{G}$ of $F(t,\mathbf{x}) - \langle F(t,\mathbf{x})\rangle$. Upon assumptions of statistical isotropy and translation invariance of the initial flame deformations one has :

$$\mathcal{G} = \frac{\Omega^2}{a^2}(S_L\Omega t)^2 g\left[\frac{X}{S_L\Omega t},\ \frac{1}{S_L\Omega t K_{neutral}},\ \frac{F_0}{S_L\Omega t},\ \frac{l_0}{S_L\Omega t}\right] \tag{31}$$

on dimensional grounds, where $X = |\mathbf{X}| = |\mathbf{x}-\mathbf{x}'|$; F_0 and l_0 represent the initial RMS amplitude of wrinkling and correlation length, respectively. The previously reported results suggest the scaling assumption

$$\mathcal{G} \longrightarrow \frac{\Omega^2}{a^2}(S_L\Omega t)^2 g\left(\frac{X}{S_L\Omega t}\right) \tag{32}$$

for $t/t_{LD} \longrightarrow +\infty$ and $X \gg \Lambda_{neutral}$. Or, if one defines the power spectrum $E(|\mathbf{k}|,t)$ by

$$\mathcal{G} = \int E(|\mathbf{k}|,t)\exp(i\mathbf{k}\cdot\mathbf{X})d\mathbf{k} \tag{33}$$

(32) is the same as

$$|\mathbf{k}|^{2+d}E(|\mathbf{k}|,t) \longrightarrow \frac{\Omega^2}{a^2}\frac{1}{\pi^{d-1}}F(\xi) \qquad \text{for } \xi \equiv \frac{|\mathbf{k}|t}{t_{LD}K_{neutral}} \ll \frac{t}{t_{LD}} \tag{34}$$

One can show [28] that $F(+\infty) \neq 0$ in (34) implies that the ensemble–averaged, effective flame speed

$$S_T \equiv S_L\left(1 + \frac{1}{2}\overline{\langle|\nabla F|^2\rangle}\right) \tag{35}$$

deduced from (14) satisfies

$$\frac{S_T}{S_L} = 1 + \frac{\Omega^2}{a^2}F(+\infty)\mathrm{Log}\frac{t}{t_{LD}} + ...,\ \frac{t}{t_{LD}} \longrightarrow +\infty \tag{36}$$

At least in the limit $\gamma \longrightarrow 0$ where $\Omega^2F(+\infty)/a^2 \longrightarrow 0$, this signals that the flame becomes fractal [29] ; the reason is that a too wide range of long ($\Lambda \sim S_Lt$ once more !) and well–developed ($t \gg t_{LD}$) wrinkles appears in the long–time limit. The fractal dimension d_F is read off from (36), viz

$$d_F \simeq d + \frac{\Omega^2}{a^2}F(+\infty) \tag{37}$$

since

$$\left(\frac{t}{t_{LD}}\right)^{d_F-d} = 1 + (d_F-d)\mathrm{Log}\frac{t}{t_{LD}} + ... \text{ as } d_F - d \longrightarrow 0 \tag{38}$$

So far, nobody could compute $F(.)$ exactly nor even analytically demonstrate that (34) holds. Accordingly (34) was checked [28] against many numerical integrations of the one–dimensional version ($d = 1$) of (14). Our findings were compatible with

$$|k|^3 E(|k|,t) \longrightarrow \frac{\Omega^2}{a^2} F(\xi) \exp\left(-\frac{|k|}{K_{neutral}}\right), \qquad F(+\infty) \lesssim 0.12 \tag{39}$$

and hence with (34) at fixed ξ. However, because of statistical bars and of forcing by numerical round-off that prevented to obtain reliable runs beyond $t = 90 t_{LD}$, the above $F(+\infty)$ cannot be entirely trusted. Also, the slow growth of S_T/S_L predicted by (36) makes it very difficult to estimate $F(+\infty)$ directly, especially for $t \leq 90 t_{LD}$. At any rate this $F(+\infty)$ yields a too small d_F compared to experiments [30] or to other, presumably noisy, numerical trials [29].

Whether premixed flames can spontaneously fractalize is one of the most intriguing theoretical problems in our field. Another one is to predict d_F (if it exceeds d !) in the presence of noise.

A final remark is due. This chapter was mostly devoted to the 1–D version of the MS equation (14). About the 2–D version, the only exact available result is that it admits solutions in the form of orthogonal patterns, $F_1(t,x_1) + F_2(t,x_2)$, where $F_1(.)$ and $F_2(.)$ evolve *independently* and are thus amenable to separate pole dynamics, in addition to the one–dimensional ones [31] ; this led some [29] to argue that $F(+\infty)|_{d=2} = 2F(+\infty)|_{d=1}$. Current attempts of ours to generalize pole–dynamics to a dynamics of lines evolving under pointwise $1/x^2$ interactions so far were unsuccessful. In fact one does not even know why the 2–D cells are so close to polygons, as if some sort of strong (yet non infinite) line–tension was acting along their sides. Presumably this is related to the fact that the crests act as nearly concentrated sinks of fluid when looked at from the fresh–gas side. Needless to stress that it is difficult to explain (see subsection 4.2, however) why $\overline{\Lambda}$ again grows like $\zeta S_L t$, not to mention the value of ζ itself.

5 Parabolæ and Noise

5.1 Parabolic MS Equation

Let $\tilde{F}(t,x)$ be a "steady" solution (i.e. with a time–independent shape) to the 1–D version of the MS equation. Write $F(t,x)$ as $\tilde{F}(t,x) + \phi(t,x)$, with a small ϕ. Then the latter satisfies

$$\frac{\partial \phi}{\partial t} + a S_L \tilde{F}_x \phi_x = S_L \Omega \left(\frac{\phi_{xx}}{K_{neutral}} + I(\phi, x) \right) \tag{40}$$

Now approximate $\tilde{F}_x$ by Cx for some positive constant C. What results is a linearized form of the "parabolic" MS equation

$$\frac{\partial \phi}{\partial t} + a S_L C x \phi_x + \frac{a S_L}{2} \phi_x^2 = S_L \Omega \left(\frac{\phi_{xx}}{K_{neutral}} + I(\phi, x) \right) \tag{41}$$

which was originally designed [32] to describe finite–amplitude disturbances of a main, parabolic flame front. It also admits a pole–decomposition [32], mostly because (14) does. Indeed consider (16) for $K/K_{neutral} \to 0$ and let $\alpha = 1$, ..., $2M$ with $M < N$ label any subset of the $2N$ poles of F_x. Then the restriction of (16) to the subset may be rewritten as

$$\frac{dx_\alpha}{dt} = \Omega S_L \left(\frac{2}{K_{neutral}} \sum_{\beta \neq \alpha}^{2M} \frac{1}{x_\beta - x_\alpha} - i\varepsilon_\alpha \right) + aS_L \tilde{F}_x(t, x_\alpha) \tag{42}$$

where $\tilde{F}_x(t,x)$ is the – analytically continuated – contribution of x_{2M+1}, ...,x_{2N} to the total front shape; x_1, ..., x_{2M}, are then the poles of a "disturbance" $\phi(t,x)$ about it. As N and K are arbitrary, envisage the double limit $N \to \infty$, $K \to 0$ and $NK \sim 1$, in a way such that $\tilde{F}_x \simeq Cx$ where x_1, ..., x_{2M} sit. The resulting from of (42) is precisely the system of pole–quations for (41). That the needed distinguished limit on N and K "locally" leading to $\tilde{F}_x \simeq Cx$ exists can be proven upon considering the "fully coalesced" steady solutions of (16), with $N = N_{max}$ (Equation (19)) for $K/K_{neutral} \to 0$. Equation (41) also possesses $2\pi/q(t)$–periodic solutions, as it can be proved upon changing the space variable from x to $z = x\exp(-aS_LCt)$. The origin of this stretching is the x–dependent convection in (41), that "redshifts" the disturbances.

Although this could be worked out in the nonlinear case [32] from the pole–method, we shall content ourselves with studying the linearized version of (41), as to illustrate the important consequences of the aforementioned stretching effects. Elementary solutions then read $\phi(t,x) = \Gamma(t)\exp(iq(t)x)$, with :

$$\frac{dq}{dt} = -aS_LCq \tag{43}$$

$$\frac{d\mathrm{Log}\Gamma}{dt} = S_L\Omega|q| \left(1 - \frac{|q|}{K_{neutral}}\right) \tag{44}$$

Due to stretching $q(t) = q(0)\exp(-aS_LCt)$, as expected from above and, as a consequence of (43) the total amplification ratio is finite

$$\frac{\Gamma(\infty)}{\Gamma(0)} = \exp\left[\frac{\Omega}{aC}|q(0)| \left(1 - \frac{|q(0)|}{2K_{neutral}}\right)\right] \tag{45}$$

contrary to the case of disturbances of flat–flames ($C = 0$). This important effect of wavelength dilation, first mentionned for flames by Zel'dovich *et al.* [33] and transposed to other growth contexts by Pelcé [34], implies that $\Gamma(\infty)$ will stay moderate if $\Gamma(0)$ is small. How small ? One may notice that

$$\frac{\Gamma(\infty)}{\Gamma(0)} = \exp\left[\frac{\Omega K_{neutral}}{2aC} - \frac{\left(|q(0)| - K_{neutral}\right)^2 \Omega}{2aCK_{neutral}}\right] \tag{46}$$

Maximum amplification occurs when $|q(0)| = K_{neutral}$, the width

$$\delta k \equiv \left(\frac{aCK_{neutral}}{\Omega}\right)^{1/2} \tag{47}$$

of the above Gaussian function of $|q(0)|$ being $o(K_{neutral})$ if $C = o(K_{neutral})$, and is $\mathcal{O}\left(\exp(\Omega K_{neutral}/2aC)\right)$ within this band. The last number is very large if $C/K_{neutral} \ll 1$, as actually assumed to obtain (41) itself : flames with wide parabolic troughs are *meta–stable*, yet *exceedingly vulnerable to disturbances*.

5.2 Noise and Cell–Size Selection

The above conclusion soon suggested [35] that flames will ultimately become sensitive to external noise, as the process of crest coalescence evoked in Sec. 5 spontaneously makes wide, nearly–parabolic and long–lived cells appear. To take up the problem an additive forcing $u(t,x)$ was added to the RHS of the 1–D version of the MS equation (14). $u(t,x)$ was selected [36] to be a sum of terms $u_n(t,x)$ of the form

$$u_n(t,x) = u'\sigma_n \cos(k_n x + x_n)\cos(k_n S_L t + z_n + \varpi_n t) \tag{48}$$

where u' is the RMS value of $u(t,x)$, k_n is a wavenumber, x_n and z_n are random phases that were kept fixed in time, and $\varpi_n \sim u'\sigma_n k_n$ are random turnover frequencies. The weights σ_n were selected for each n so as to ensure that in the continuum limit the spectral energy density of forcing was proportional to $[u']^2 S(|k|/K_{int})/K_{int}$, where K_{int} is an integral wavenumber and $S(.)$ specifies the spectrum shape.

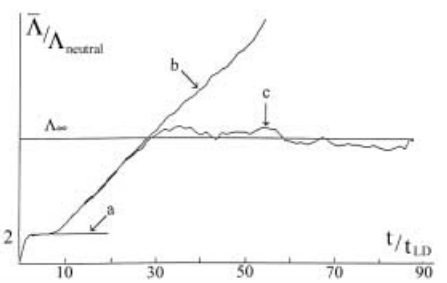

Fig. 14. Mean, ensemble–averaged, cell wavelengths $\overline{\Lambda}(t)$ vs time. (a) forced, linearised MS equation ; (b) nonlinear MS equation, with forcing suppressed for $t > 8t_{LD}$; (c) nonlinear MS equation, forced from $t = 0$ on

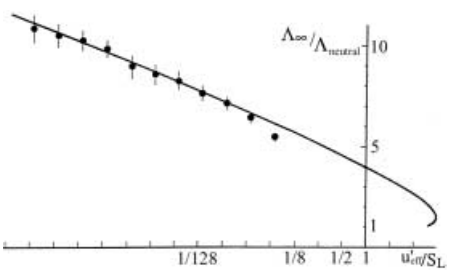

Fig. 15. Λ_∞ versus the effective noise intensity u'_{eff} (Eq. (51)). *Symbols* : numerical results ; *Solid line* : from Eq. (51)

The typical ensemble–averaged results of numerical integrations of the 1–D, forced MS dynamics can be summarized as follows [36] [37].

1— For $u'/S_L \ll 1$ the mean spacing $\overline{\Lambda}(t)$ between front crests starts evolving as if forcing was *cut off* after a few t_{LD} (cf. Equation (20)). In particular $d\overline{\Lambda}(t)/dt = \zeta S_L \Omega$ for a long while, with the same ζ as in subsection 4.1 (see Fig. 14).

2— Rather abruptly $\overline{\Lambda}(t)$ saturates to a value Λ_∞ which for a given spectral function $S(.)$ and a fixed $K_{int}/K_{neutral}$ increases slowly ($\sim$ logarithmically) as $u'\Omega/aS_L$ decreases ; Fig. 15. The existence of Λ_∞ results from *crest implants* in the largest cells.

3— For a fixed $u'\Omega/aS_L$ and $K_{int} = K_{neutral}$, $\Lambda_\infty K_{neutral}$ does not depend on $S(.)$ significantly (only $\sim$ logarithmically). Λ_∞ depends only mildly on $K_{int}/K_{neutral}$ and on $S(.)$, at least whenever the latter varies algebraically for large/small arguments.

A tentative (and admittedly crude) theoretical explanation [36] [37] of the above properties, based on ideas forwarded in [33] [35], is summarized below. It rests on the following ideas.

• In the asymptotic regime where $\overline{\Lambda}(t) = \Lambda_\infty$, each cell may on average be approximated by a parabola, at least in some vicinity of its trough and if $\Lambda_\infty/\Lambda \gg 1$.

• This average parabolic cell cannot be too wide, for otherwise the noise would induce secondary wrinkles that are large enough to break it, and it could not be identified as the mean cell any longer.

For a parabolic front of curvature C at its trough, the maximum amplification ratio $\Gamma(\infty)/\Gamma(0)$ of subwrinkles that had $|q(0)|$ as initial wavenumber is given by (46), and hence is $\sim \exp(\Omega K_{neutral}/2aC)$ when $|q(0)|$ lies in the wavenumber band of width δk defined in (47). Following a suggestion originally made by Clavin [35], one can next estimate $\Gamma(0)$ dimensionally as $\Gamma(0) \sim \delta u'.\delta t$, where $\delta t \sim \Lambda_{neutral}/S_L$ is the time of interaction between the flame and the most dangerous modes of forcing ($k_n \simeq K_{neutral}$ in (48)). Indeed each contribution $u_n(t,x)$ to $u(t,x)$ given by (48) corresponds to flame propagation at speed $\sim S_L$ through a chessboard array of spatially $2\pi/k_n$–periodic square eddies. As for $\delta u'$, we estimate it from the energy of forcing in the "dangerous" band of δk width, centered about $K_{neutral}$, viz :

$$[\delta u']^2 \sim \frac{\delta k}{K_{int}}[u']^2 S\left(\frac{K_{neutral}}{K_{int}}\right) \tag{49}$$

Putting all the pieces together yields

$$\Gamma(\infty)K_{neutral} \sim \frac{u'}{S_L}\left(\frac{aC}{\Omega K_{neutral}}\right)^{1/4}(\lambda S(\lambda))^{1/2}\exp\left(\frac{\Omega K_{neutral}}{2aC}\right) \tag{50}$$

where $\lambda \equiv K_{neutral}/K_{int}$

Finally C is estimated as $2\Omega/a\overline{\Lambda}$, i.e. like in large steady crests [21], and the mean wrinkle amplitude $\overline{A}$ as $\sim \overline{\Lambda}\Omega/a$. One can now apply the proposed *selection criterion* : Λ_∞ is such that $\Gamma(\infty) \sim \overline{A}$. This results in the equations

$$\xi^{5/4}\mathrm{e}^{-\frac{\pi}{2}\xi} \sim \frac{a}{\Omega}.\frac{u'}{S_L}(\lambda S(\lambda))^{1/2} \equiv \frac{u'_{eff}}{S_L} \tag{51}$$

$$\xi \equiv \frac{\Lambda_\infty}{\Lambda_{neutral}}\ , \qquad \lambda \equiv \frac{K_{neutral}}{K_{int}} \tag{52}$$

that give [2] ξ as a function of the main quantities characterizing the mixture and the noise properties. One may check that (51) (52) *captures* all the properties 1) — 3) evidenced by numerical simulations of the forced MS equation, lending some credence that the above theory focuses on the good mechanisms. Putting it in a more rigorous form has not yet been achieved, though pole dynamics combined with pole–sparkling [38] to simulate the noise hopefully will serve as helpful starting points [39].

A last remark on the matter is in order. The exponential relationship between u'/S_L and $\Lambda_\infty/\Lambda_{neutral}$ implies that even round–off numerical noise is able to quench cell–coalescence ; e.g. working with 14-digit numbers yields $\Lambda_\infty \simeq 26\Lambda_{neutral}$ [37]. In Nature, even molecular fluctuations in composition cannot [40] be neglected ultimately. The same kind of remark applies to flame propagating in spray/air mists [40], though at a stronger noise level.

6 Expanding Fronts

We so far focused our attention on flat–on–average fronts, the simplest (?) theoretical case but the most difficult configuration to study experimentally [41]. A more convenient experimental procedure is to ignite locally the mixture, thereby triggering an outwardly expanding front. How this overall growth in size is to date known to affect the front dynamics, is summarized in this Sec. 6.

6.1 Expansion and Landau–Darrieus Instability

When a mixture is ignited by a wire (or a spark), idealized as a line (or a point) located at $r = 0$, a circular (or spherical) flame of radius $r = R(t)$ expands. If all curvature/stretch effects are neglected (i.e. $S_n \equiv S_L$) mass conservation requires

$$\frac{dR}{dt} = S_b \equiv \frac{S_L}{1-\gamma} \tag{53}$$

since the compressible burnt gases stay at rest in the flame center frame. Outside the flame, the radial flow velocity is $S_L(R/r)^2\gamma/(1-\gamma)$. In less ideal situations the front is a *deformed* circle (or sphere) of equation $r - S_b t \sim \Gamma(t)\mathrm{e}^{in\theta}$ (or $\sim \Gamma(t)\mathrm{e}^{im\phi}P_n^m(\cos(\theta))$, P_n^m being the associated Legendre function) in cylindrical (or spherical) coordinates (r,θ) (or (r,θ,ϕ)). Assuming $\Gamma/R \ll 1$, solving the linearized Euler + Hugoniot equations ultimately yields

$$\frac{d\,\mathrm{Log}\Gamma}{dt} = S_L\frac{h(\gamma,|n|)}{R(t)} \tag{54}$$

in both cases, where $h(.,.) \geq 0$ satisfies $h(0,|n|) = 0$, $h(\gamma,0) = h(\gamma,1) = 0$, $h(\gamma,|n| \gg 1) = |n|\Omega(\gamma)$. The wrinkle growth rate must indeed vanish when

[2] An equation for $\mu \equiv aC/\Omega K_{neutral}$ could have been written as an alternative to (51), viz $\mu^{-5/4}e^{-1/\mu} \sim u'_{eff}/S_L$

either $\rho_u = \rho_b$ (no Landau–Darrieus mechanism), $n = 0$ (a mere shift in time), $|n| = 1$ (a shift of the whole flame), and ought to reduce to (3) with $|k| = n/R(t)$ in the limit of large harmonic numbers. For $|n| \gg 1$ and $\gamma \ll 1$ one thus has $h(\gamma, |n|) = |n|\gamma/2 + ...$ The exact expression of $h(\gamma, |n|)$ in the spherical case can be found in [42]. Because $R = S_b t$, the wrinkle amplitude grows *algebraically* in time, $\Gamma \sim t^{h(\gamma,|n|)(1-\gamma)}$, as opposed to exponentially in the case of nearly–planar fronts.

6.2 MS Equations for Deformed Circles or Spherical Sectors

For $\gamma \ll 1$, one can adapt the analysis giving (12) to a nearly–circular front of equation $r = R(t) - F(t, \theta)$, with here $\langle F(t, \theta) \rangle \equiv 0$ — $F(t, \theta)$ is then the front deformation — assuming that the current cell wavelengths are $\mathcal{O}(\Lambda_{neutral})$ which implies $|n| \sim 1/\gamma \gg 1$. The resulting "cylindrical" MS equation [31] [43] [44] can be written as :

$$\frac{\partial F}{\partial t} + \frac{a S_L}{2R^2} \left(F_\theta^2 - \langle F_\theta \rangle^2 \right) = S_L \Omega(\gamma) \left(\frac{F_{\theta\theta}}{R^2 K_{neutral}} + \frac{I(F, \theta)}{R} \right) \tag{55}$$

where $\langle . \rangle$ and $I(., \theta)$ ($\equiv$ multiplication by $|n|$ in the angular Fourier space) now refers to the angle. Although eq. (55) has to date been derived only to leading order in γ ($\Omega(\gamma) = \gamma/2 + ...$) the equation for $F(t, \theta)$ was written in the above form because it reduces to (14) (up to a trivial change of unknown) in the double limit $RK_{neutral} \longrightarrow \infty$, $R\theta K_{neutral} = \mathcal{O}(1)$ and hence is believed to yield fair quantitative results even when γ is not small. The linearized version, once Fourier–transformed, leads to

$$\frac{d \, \mathrm{Log} \Gamma}{dt} = S_L \Omega \left(\frac{|n|}{R} - \frac{n^2}{R^2 K_{neutral}} \right) \tag{56}$$

to be compared to (7) and (44); (56) is well confirmed by direct numerical simulations [48]. As for $R(t)$, it should be determined from

$$\frac{dR}{dt} = S_b \left(1 + \frac{1}{2} \langle F_\theta \rangle^2 \right) + \text{curvature term.} \tag{57}$$

However, we shall simply take $R = S_b t$ here, because this has no identified qualitative impact on the results to be summarized next. The case of *equatorial sectors* (see Fig. 16) of deformed spheres can be handled similarly. The equation for $F(t, \vartheta, \phi)$, where $\vartheta = \pi/2 - \theta$ is the latitude, is obtained from (55) via the substitutions : $F_\theta^2 \to F_\vartheta^2 + F_\phi^2$, $F_{\theta\theta} \to F_{\vartheta\vartheta} + F_{\phi\phi}$ and $I(., \theta) \to I(., \vartheta, \phi)$ (multiplication by $(p^2 + m^2)^{1/2}$, called again $|n|$, in the (p, m) 2–D angular Fourier [3] space).

Numerical integrations of (55) and its 2–D generalisation [45] [46] revealed the following : random deformations at initial at time t_0 ultimately evolve into a

[3] Far for the sphere poles, Surface Harmonics go to trigonometric functions in the limit of large orders and degrees.

Fig. 16. Snapshots of equatorial sectors of 3–D expanding flames, as computed from the 2–D version of (55). *Left* : spontaneous evolutions from random initial deformations ; *Right* : emergence of small subcells (with $\overline{\Lambda} \sim 5\Lambda_{neutral}$) induced by a small additive forcing

few localized and angularly fixed crests (Fig. 16 left), whose number is smaller if the initial radius is small, even if t noticeably exceeds the characteristic growth time t_{LD} defined in (20).

This is in qualitative agreement with the experimentally observed (yet not-too-late) dynamics [42] [47]. This is rather astonishing : why don't wrinkles grow on the increasingly large flame front smooth portions in between the localized crests ? Also, why do only a few crests survive ? Theoretical attempts of explanation are greatly eased by the fact that (48) also admits a pole-decomposition [31] [43], now in terms of the complex angular singularities $\vartheta_\alpha(t)$ of $F_\theta(t,\vartheta)$. We merely quote below the system of ODEs analogous to (16) :

$$\frac{d\vartheta_\alpha}{dt} = \Omega S_L \left[\frac{1}{R^2 K_{neutral}} \sum_{\beta \neq \alpha}^{2N} \operatorname{cotan}\left(\frac{\vartheta_\beta - \vartheta_\alpha}{2} \right) - i \frac{\varepsilon_\alpha}{R} \right] \tag{58}$$

the corresponding $F(t,\vartheta)$ being expressed as a sum of $Log\,(sin(\vartheta - \vartheta_\alpha)/2)$ functions , like in (15).

Like in (18) it is again revealing to analyse the "collision" of two nearby poles ϑ_1 and ϑ_2. For $\vartheta_1 \simeq \vartheta_2$, (58) yields

$$\frac{d(\vartheta_1 - \vartheta_2)}{dt} \simeq \frac{4\Omega S_L}{R^2 K_{neutral}} \frac{1}{\vartheta_2 - \vartheta_1} \tag{59}$$

and, once more, the two poles tend to align parallely to the imaginary ϑ–axis. The final alignment could be quite imperfect, however, as $\int^\infty dt/R^2 < \infty$ if $R \geq S_b t$. Besides, because $\int^\infty dt/R = \infty$, all the ϑ_α's must ultimately approach the real axis ($\mathcal{I}\,(\vartheta_\alpha(t)) \sim 1/RK_{neutral}$). These properties explain qualitatively why crest coalescence is very efficient during the early flame growth ($R^2 K_{neutral}$ is not large then), "removing" most of the initial front wrinkles, but is quenched by expansion at later times (it makes the crests move apart). Before embarking on a semi–quantitative description of the process, a remark is due. Even though all the initially present poles must ultimately "land" on the real ϑ–axis, only a limited number of them do [31] during reasonable times ($t \leq 30 t_{LD}$, say).

The reason is similar to that leading to the existence of N_{max} in the "planar" case (cf. equation (19)). Once enough of pole pairs have already landed, the repulsive interactions encoded in (58) prevent the other ones from doing so soon. Furthermore $\int^t d\tau/S_b\tau$ diverges only slowly for $t \longrightarrow \infty$.

6.3 Another Mean-Field Theory, for Deformed Circles

Being explicitly time-dependent the pole-dynamics (58) associated with the MS equation for deformed circles is even more difficult to study analytically than for flat-on-average fronts. To try describe the coalescence/expansion in statistical terms we developed [45] again a mean-field approach, similar in spirit to the Lifchitz-Slyozov-Wagner (LSW) theory for the statistics of precipitated grains in supersaturated media [49]. The LSW analysis first focuses on a typical, anonymous grain of size a and determines $da/dt = V(a, \Delta(t))$ as a function of a and instantaneous **mean** supersaturation $\Delta(t)$. A kinetic equation $\partial f/\partial t + \partial(fV)/\partial a = 0$ then describes the function $f(a,t)$ such that $daf(a,t)$ is the mean number density of grains with sizes in $[a, a+da]$ at time t. The subtle point is that such an hyperbolic equation is constraint by an overall conservation law, namely that of the solute, that relates $\Delta(t)$ to some moment of $f(a,t)$. We proceeded along similar lines. Considering a collection of crests located at angles $\Theta_1, \cdots \Theta_{N_c}$ the zeroth step was to show that, if N_c and $2\pi R/\Lambda_{neutral}$ are large enough, one may indeed reduce (58) to a restricted N_c–body problem similar to (23) but now along the circle, for the Θ_ns. To study the latter, the second step was to focus on one anonymous cell of angular size ϑ ($= \Theta_{n+1} - \Theta_n$ for some n) ; upon approximating the influence of "the other crests" like in paragraph 4.3, the approximate law

$$\frac{d\vartheta}{dt} \cong V(\vartheta, t) = \frac{4\Omega S_L}{R^2 K_{neutral}} \frac{P_n + P_{n+1}}{2} \left(-\frac{1}{\vartheta} + \frac{\vartheta}{\vartheta_{av}^2(t)} \right) \tag{60}$$

was found ; though very similar to (25), it differs slightly as the unknown quantity $\vartheta_{av}(t)$ is *not* exactly the mean angular cell-size $\overline{\vartheta}(t)$, for reasons explained below ; notice that $\vartheta_{av}(t)$ plays the same part as the critical grain radius $a_{crit}(t)$ did in the LSW theory : cells with $\vartheta > \vartheta_{av}$ (resp. $\vartheta < \vartheta_{av}$) tend to spontaneously grow (resp. shrink). The third step was to identify the mean weight $(P_n + P_{n+1})/2$ of the "locally" interacting crests with the average weight $P_{av}(t)$ like in 4.3. The fourth step consisted, like in the LSW theory , in writing the kinetic equation

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \vartheta}(fV) = 0 \tag{61}$$

where $d\vartheta f(\vartheta, t)$ is the mean number of cells with angular sizes in $[\vartheta, \vartheta + d\vartheta]$. Now the *overall conservation laws*. There are two of them, namely

$$P_{av}(t) N_c(t) = P_0 N_0, \qquad \text{with} \qquad N_c(t) \equiv \int\limits_{\vartheta > 0} f(\vartheta, t) d\vartheta \tag{62}$$

and :

$$\int\limits_{\vartheta>0} \vartheta f(\vartheta,t)d\vartheta = 2\pi \tag{63}$$

The first one expresses that coalescences are mere pole rearrangements, whereas the second ensures that all cell angular sizes must sum up to a full turn [60]. One can show that (63) implies

$$\vartheta_{av}^2(t) = \int\limits_{\vartheta>0} \vartheta f(\vartheta,t)d\vartheta / \int\limits_{\vartheta>0} f(\vartheta,t)d\vartheta/\vartheta \tag{64}$$

so that in general $\vartheta_{av} \neq \overline{\vartheta} \equiv \int \vartheta f(\vartheta,t)d\vartheta / \int f(\vartheta,t)d\vartheta = 2\pi/N_c$; actually $\vartheta_{av} \leq \overline{\vartheta}$. Using $\overline{\vartheta}$ instead of ϑ_{av} in (60) leads to the cropping up of many spurious small cells to "fill the gap", as (63) is not satisfied. Fortunately (60) – (64) could be solved analytically [45] leading to results that are in fair agreement (Fig. 17) with those of numerical simulations. One may note that expansion always wins over coalescence, leaving over a finite number $N_c(\infty) \leq N_0$ of crests ; in some cases, coalescence is totally suppressed, e.g. when N_0 or $1/t_0$ are too small, and/or if $f(\vartheta,t)$ is too peaked about its mean.

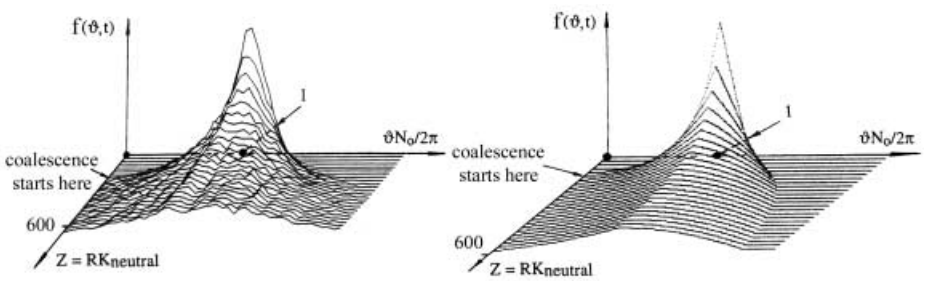

Fig. 17. Evolutions of cell angular–size distribution function $f(\vartheta,t)$. *Left* : from ensemble–averaged numerical integrations of (55) ; *Right* : from the kinetic equation (61). Both have $R = S_L t/(1-\gamma)$ and a triangular $f(\vartheta,t_0)$

6.4 Expanding Flames and Noise

While qualitatively reproducing the not-too-late histories of experimental flames, (55) and/or its 2–D generalisation are apparently unable to mimic the late and sudden cropping up of many subcells between the few isolated crests created by the initial conditions, almost simultaneously all over the front. A tentative explanation [46] [50], that invokes noise (for a change !), is summarized below. To this end consider the linearised, then Fourier transformed ($F(t,\theta) \leftrightarrow \Phi_n(t)$)

equation (55) when an additive noise $u(t,\vartheta)$ is added to the RHS. The Fourier components $\Phi_n(t)$ of $F(t,\vartheta)$ satisfy

$$\frac{d\Phi_n}{dt} = \Omega S_L \left(\frac{|n|}{R} - \frac{n^2}{R^2 K_{neutral}} \right) \Phi_n + u_n(t) \tag{65}$$

The forcing $u(t,\vartheta)$ is assumed broad-banded at all times and, for simplicity, white in time, so that [4] $\overline{u_n(t)u_m(t')} = D(m,t)\delta_{m+n}\delta(t-t')$; $D(m,t) \geq 0$ specifies the angular spectrum of $u(t,\vartheta)$. It is a simple matter to compute $\Phi_n(t)$ then $\overline{|\Phi_n(t)|^2}$ which, up to an unimportant prefactor, is found for $t/t_{LD} \to \infty$ to read as

$$\overline{|\Phi_n(t)|^2} \sim tD\left(|k|R, \frac{t|k|}{K_{neutral}}\right) e^{-\Omega S_L t\left(|k|\left(1-\frac{|k|}{K_{neutral}}\right)+|k|Log\frac{|k|}{K_{neutral}}\right)} \tag{66}$$

where $|k| \equiv |n|/R(t)$ is assumed fixed to be less than $K_{neutral}$ (otherwise $\overline{|\Phi_n|^2} \sim DR^2/n^2$). For $t/t_{LD} \to \infty$, $\overline{|\Phi_n|^2}$ is strongly peaked about $|k| = k_c \equiv \kappa K_{neutral}$, where $\kappa \simeq 1/4.92 < 1$ is the smaller root of $Log\kappa + 2(1-\kappa) = 0$. The peak half-width is $\mathcal{O}\left(\sqrt{t_{LD}/t}\right)$ and, more importantly, its maximum is $\sim tD\left(k_c R, tk_c/K_{neutral}\right) e^{8\kappa(1-\kappa)t/t_{LD}}$. Consequently

$$\left(\overline{F(t,\theta)^2}\right)^{1/2} \sim \left(tD\left(k_c R, \frac{tk_c}{K_{neutral}}\right)\right)^{1/2} \left(\frac{t_{LD}}{t}\right)^{1/4} e^{4\kappa(1-\kappa)\frac{t}{t_{LD}}} \tag{67}$$

and the noise-induced wrinkles pass from hardly noticeable to clearly visible for $t = t_* + \mathcal{O}(t_{LD})$, where

$$\left(\frac{t_{LD}}{t_*}\right)^{1/4} e^{4\kappa(1-\kappa)\frac{t_*}{t_{LD}}} \sqrt{t_* D} K_{neutral} \sim \frac{\Omega}{a} \tag{68}$$

beyond which time nonlinearity comes into play. All the newly born cells then have about the same wavelength $\overline{\Lambda} \simeq 5\Lambda_{neutral}$ (instead of $2\Lambda_{neutral}$ in the planar case) and, because of the exponential featured in (68), even the tiniest forcing will ultimately produce this phenomenon. One must also notice that only a small region of the exciting field matters ; if $D(\cdot,\cdot)$ is, e.g., of the form $D(m/K_{int}R)$ and hence is characterized by a single integral wavenumber K_{int} from $t=0$ on, it is $D(k_c/K_{int})$ which enters (67) (68).

Numerical integrations of (55) and of its 2–D generalization fully confirmed [46] the above predictions on the noise-induced wrinkling, even in presence of a few isolated crests caused by initial distorsions ! see Figure 16 (*Right*). The theoretically predicted sequence of events also is in agreement with what is observed in the laboratory [42] [47], suggesting that some noise is acting : residual turbulence, molecular composition noise, dust ... feed-back of vortical interactions in the burnt gases on the fresh gas motions ...

What happens for $t \gg t_*$ is still unclear, yet one may surmise that noise, whatever this would represent, will ultimately make the mean cell wavelength

[4] $\overline{(\cdot)}$ means: ensemble average ; δ_{m+n} is the Kronecker delta

saturate at a finite value Λ_∞, like in the planar-on-average situation. Indeed, at the scale of Λ_∞ a growing sphere looks like a plane. One has to realize, however, that 4.92 is not a small number : only very low levels of forcing would yield $\Lambda_\infty > 4.92\Lambda_{neutral}$, assuming that (51) still holds ! The possibility that the mean crest-to-crest distance $\overline{\Lambda}$ would saturate at a finite value Λ_∞ also raises an intriguing question : is the existence of Λ_∞ compatible with a front that evolves towards a fractal ?

6.5 Model Equations for Whole Deformed Spheres

The 2–D generalization of (55) has virtues and disavantages : it can be handled by means of Fourier pseudo-spectral integration methods, but can only describe an equatorial sector ; this limitation introduces also some degree of arbitrariness, via the boundary conditions that are needed along the sector's side. To circumvent the difficulty we recently introduced [50] a class of model evolution equations for that : (i) describe $F(t,\theta,\phi)$ over the whole deformed sphere ; (ii) are invariant against finite rotations and small translations of the coordinate system ; (iii) reduce to (14) and to the 2–D version of (55) for $RK_{neutral} \gg 1$ and suitably-selected angular ranges ; (iv) be asymptotically correct for $\gamma \to 0$ uniformly in (θ,ϕ), the sphere poles included. In a sense, such model evolution equations essentially incorporate all what we summarized before in this essay. They read as

$$\frac{\partial F}{\partial t} \;=\; \Omega S_L \left(\frac{\nabla_S^2 F + 2F}{R^2 K_{neutral}} + \frac{H(F,\theta,\phi)}{R} \right) \;+\; \frac{a(\gamma) S_L}{2R^2} \left| \nabla_S F \right|^2 \;+\; CT \qquad (69)$$

where ∇_S^2 is the surface Laplacian (hence $\nabla_S^2 + 2$ accounts for local curvature), $|\nabla_S F|^2$ is the squared surface gradient. The linear hydrodynamic operator $H(F,\theta,\phi)$ generalises $I(F,\vartheta)$ and is defined by : $H(e^{im\phi} P_n^m(cos\theta),\theta,\phi) \equiv \frac{2n(n-1)}{2n+1} e^{im\phi} P_n^m(cos\theta)$. Because ∇_S^2 is also diagonal in the basis of surface harmonics (its symbol there is $-n(n+1)$), (69) is well-suited for Fourier-Legendre pseudo-spectral integrations. Finally CT is a counterterm, that kills the $n=0$ $(m=0)$ and $n=1$ $(m=0,1)$ modes of $F(t,\theta,\phi)$, for those do not correspond to genuine deformations. Figure (18) shows sample results obtained numerically [50] with or without additive forcings. The visual agreement with experiments on "free" [42] [47] or weakly-turbulent [51] [52] flames is striking. Understanding the long-time front dynamics in presence of noise is one of the most challenging, yet unsolved, problems of flame theory.

7 Concluding Remarks. Open Problems

Even though flame-front theory entered the nonlinear stage quite a while ago [12] many problems, ranging from physical to purely mathematical, remain unraveled. A few of them, of a more statistical flavor are briefly evoked below.

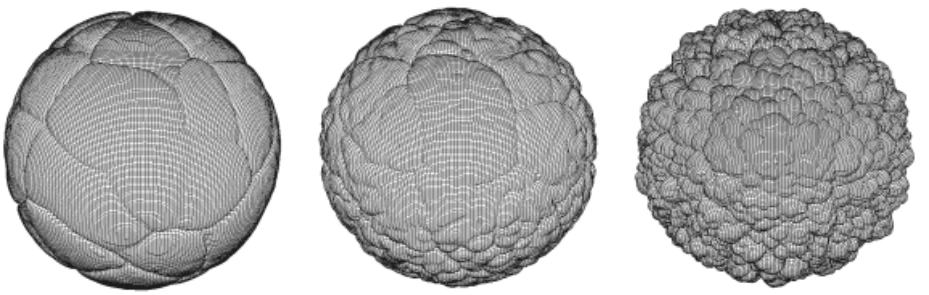

Fig. 18. Snapshots of 3–D expanding flames as predicted from Eq. (69). *Left* : spontaneous evolution, random initial deformations ; *Middle* : emergence of small subcells (with $\overline{\Lambda} \sim 5\Lambda_{neutral}$) induced by a small additive forcing ; *Right* : "pebbled" flame front resulting from a stronger forcing

7.1 1–D Flame Fronts

In the context of such 1–dimensional MS–type of equations as (14) (41) (55), the coherent structures we identified along the fronts are its cell crests. Elucidating how those statistically evolve from initial conditions is a central, only partly understood, problem. A first step could be to exactly solve the statistical problems associated with the N–body problem (23) and its generalisation to expanding flames. One has not to forget, however, that crests are mere "shadows" on the real axis of a movie whose stage lies in the complex plane. Developing methods to theoretically handle the statistics of pole motions would be even more useful. Of course, finding the general solutions to (14) (41) (55) is a Holly Graal ! Indeed, exhibiting an infinite number of solitons of a nonlinear PDE often is insufficient for this purpose (see [23] for examples). Handling the influence of noise in a clean way would also be of tantamount importance, both theoretically and practically, as most actual flames encounter some turbulence, even if only weak. Putting the analyses of Sec. 5 into a rigorous framework certainly is a prerequisite.

One must also stress that the sections devoted to nearly–planar flames focused on temporally–developing patterns, i.e. on their time–wise evolution from initial conditions at fixed transverse locations. Though theoretically and computationally convenient this formulation is rather impractical for comparisons with experiments, except in carefully controlled circumstances [41]. The situation where one imposes a significant flow–velocity component $U_{\parallel} \gtrsim \mathcal{O}(S_L)$ parallely to the mean front (Fig. 19) is much more convenient experimentally [53] and often encountered in practice. A simple equation to model $G(t,x) \equiv \partial F/\partial x$ in this case reads

$$\frac{\partial G}{\partial t} + U_{\parallel}\frac{\partial G}{\partial x} + S_L a(\gamma) G \frac{\partial G}{\partial x} \;=\; S_L \Omega(\gamma)\left(\frac{1}{K_{neutral}}\frac{\partial^2 G}{\partial x^2} + I(G,x)\right) \qquad (70)$$

in analogy with the x–differentiated, 1–D MS equation (14). A forcing term

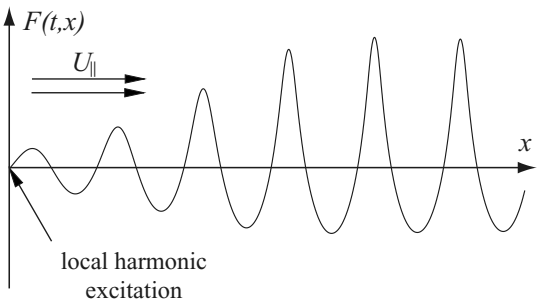

Fig. 19. Spatially developing flame instability triggered by a local harmonic excitation in presence of a large tangential flow velocity $U_{\|} \gg S_L$

$h(t)d\delta(x)/dx$ could be added to the RHS of (70) to mimic a local forcing (like in the experiments of Ref. [53]). Interestingly enough, the limit $U_{\|}/S_L \to \infty$ converts (70) into an ordinary MS equation where $\tau \equiv x/U_{\|}$ and $\chi \equiv x - U_{\|}t$ play the parts of time and space, respectively ; at least, this holds [54] when $h(t)$ is $2\pi/\omega_e$–periodic and ω_e is chosen to yield $K \equiv \omega_e/U_{\|} = \mathcal{O}(K_{neutral})$. For $1/3 < K/K_{neutral} < 1$ the 2–pole $2\pi/K$–periodic solutions (15), once written in terms of τ and χ, encountered some success when compared to experiments where $h(t) \sim h\cos(\omega_e t)$, $hK_{neutral}a/\Omega \ll 1$ [54]. Now an open question : when $h(t)$ is random, will the mean flame–brush thickness associated with (70) increase with x like $\zeta\Omega(\gamma)S_L x/U_{\|}$, ζ being the same as in Section 4 ? An analogy with the Kelvin–Helmoltz instability (again $\omega \sim |k|$ at small $|k|$!), vortex pairing being meant as an analogue of crest–merging, would then be worth exploring.

7.2 2–D Flame Fronts

Two–dimensional, flat–on–average, flame fronts spontaneously evolving from random initial deformations ultimately experience cell-coalescence (Fig. 20). Our numerical findings (Fig. 7) are compatible with $\overline{\Lambda} \sim S_L\Omega t$ in the long–time limit, yet weaker (e.g. logarithmic) multiplicative corrections cannot be excluded. Despite the similarity with 1–D situations, such evolutions raise many unsolved theoretical problems. For 1–D evolutions the pole–dynamics naturally led to focus on the front crests, while the troughs were actually as numerous and hence could have served as equivalent starting points for statistical studies. When two transverse coordinates are involved, crests and troughs no longer play equivalent roles ; the answer to the question "why are the cell sides so close to straight lines in Fig. 20 ?" might well reside in the remark that crests are mere borders between cells. After all, the crests are just where neighbouring fronts collide, whereas propagation itself mainly occurs over the whole cells centered about the front troughs. The model equation (re–written with our notation) studied analytically in [52]

$$\frac{\partial F}{\partial t} + aS_L\left|\nabla F\right|^2 \;=\; S_L\Omega(\gamma)\overline{k}F \qquad , \qquad 0 < \overline{k} = const \tag{71}$$

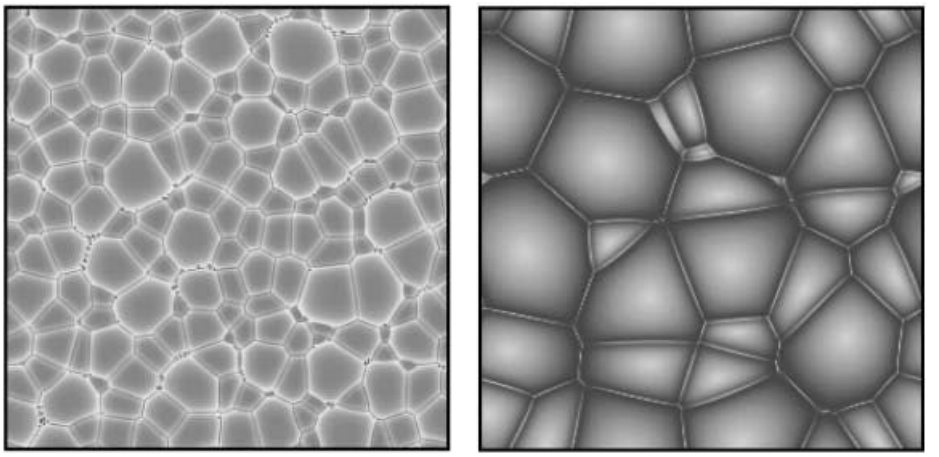

Fig. 20. Coarsening cells of 2–D, plane on average, flame fronts (time runs from left to right) ; courtesy of C. Josserand [59]

leads to cell shapes/evolutions which look very similar to what Fig. 20 and experiments show. While this militates in favor of the conjecture that the cells (as opposed to the crests) are *the* good objects to look at irrespective of the involved number of transverse coordinate(s), one must not forget the following : plane–filling 2–D polygonal patterns are so tightly constrained by topology [55] that they tend to all look alike when eye–balled. It is now patent that our understanding of flame evolutions like in Fig. 20 is pretty poor : tools and/or concepts taylored by the statistical–physics community will certainly be invaluable.

7.3 Last Points

The above essay is incomplete, possibly not impartial, in many aspects : e.g. we said nothing on the influence of body–forces (such as gravity) upon flame behaviors, especially on the very important acoustic/combustion couplings. Simply, nothing significantly nonlinear is known on the matter. To counterbalance that and have access to other facets of combustion dynamics not covered above, the reader is kindly urged to consult [24] [41] [56] [58] and the references therein.

Acknowledgements

We do thank the organizers of the $XVII^{th}$ Euroconference for an invited Lecture, hospitality at Sitges seashore and Palau Maricel,... and patience about the proceedings. Also, they allowed us to present more detailed results on subsection 6.5 at poster session [57] of the meeting. Thanks again.

References

1. Clavin, P., Joulin, G. (1983) *Premixed flames in large scale and high intensity turbulent flow* J. Phys. Lett. (France) **44**-1, L1
2. Garcia-Ybarra, P.L., Nicoli, C., Clavin, P., (1984), , Combust. Sci. Tech., **40**, 41
3. Deshaies, B., Cambray, P., (1990), *The velocity of a premixed flame as a function of the flame stretch : an experimental study*, Combust. Flame, **82**, 361
4. Quinard, J., (1984), *Limites de stabilité et structures cellulaires dans les flammes prémélangées – Etude expérimentale*, Thèse de Doctorat, Université de Provence, Marseille (France)
5. Markstein, G.H., (1964), *Nonsteady flame propagation*, Oxford : Pergamon Press
6. Joulin, G., Vidal, P., (1998), *An introduction to the instabilities of flames, shocks and detonations*, in Hydrodynamics and nonlinear instabilities, Godrèche and Manneville (Eds), Cambridge University Press, 493
7. Landau, L.D., (1944), *On the theory of slow combustion*, Acta Physicochimica URSS, **19**, 77
8. Darrieus, G., (1938), *Propagation d'un front de flamme* Unpublished work presented at La Technique Moderne, Paris (France)
9. Pelcé, P., Clavin, P., (1982), *Influence of hydrodynamics and diffusion upon the stability limit of laminar premixed flames*, J. Fluid Mech., **124**, 219
10. Frankel, M., Sivashinsky, G.I., (1982), *The effect of viscosity on hydrodynamic stability of a plane flame front*, Combust. Sci. Tech., **29**, 207
11. Garcia-Ybarra, P.L., Clavin, P., (1984), , J. Mécanique Théor. et Appl. (France), **2**, 245
12. Sivashinsky, G.I., (1977), *Nonlinear analysis of hydrodynamic instability in laminar flames. Part 1 : derivation of basic equations*, Acta Astronautica, **4**, 1177
13. Sivashinsky, G.I., Clavin, P., (1987), *On the nonlinear theory of hydrodynamic instability in flames*, J. Phys. (France), **48**, 193
14. Joulin, G., Cambray, P., (1992), *On a tentative, approximate evolution equation for markedly wrinkled premixed flames*, Combust. Sci. Tech., **81**, 243
15. Boury, G., Joulin, G., (2000), *On perturbation approaches to the nonlinear dynamics of premixed wrinkled flames*, Combust. Sci. Tech., Submitted
16. Bychkov, V.V., Golberg, S.M., Liberman, M.A., Eriksson, L.E., (1996), *Propagation of curved stationary flames in tubes*, Phys. Rev. E, **54**-4, 3713
17. Bychkov, V.V., (1998), *Nonlinear equation for curved stationary flame and the flame velocity*, Phys. of Fluids, **10**-8, 2091
18. Bychkov, V.V., Kovalev, K., Liberman, M., (1999), *Nonlinear equation for curved nonstationary flames and flame stability*, Phys. Rev. E, **60**-3, 2897
19. Denet, B., (1988), *Simulation numériques d'instabilités de fronts de flamme*, Thèse de Doctorat, Université de Provence, Marseille (France)
20. Kadowaki, S., (1999) *The influence of hydrodynamic instability on the structure of cellular flames*, Phys. of Fluids, **11**-11, 1
21. Thual, O., Frisch, U., Henon, M., (1985), *Application of pole-decomposition to an equation governing the dynamics of wrinkled flame fronts*, J. Phys. (France), **46**, 1485
22. Lee, Y.C., Chen, H.H., (1982), *Nonlinear dynamical models of plasma turbulence*, Physica Scripta, **2**, 41
23. Ablowitz, M.J., Clarkson, P.A., (1991), *Solitons, nonlinear evolution equations and inverse scattering*, Cambridge University Press

24. Vaynblat, D., Matalon, M., (2000) *Stability of pole solutions for planar propagating flames : I. exact eigenvalues and eigenfunctions*, SIAM J. Applied Math., in press
25. Cambray, P., Joulain, K., Joulin, G., (1994), *Mean evolution of wrinkle wavelengths in a model of weakly-turbulent premixed flame*, Combust. Sci. Tech., **103**, 265
26. Bessis, D., Fournier, J.D., (1984), *Pole condensation and the Riemann surface associated with a shock in Burgers' equation*, J. Phys. Lett. (France), **45**, L833
27. Frisch, U., (1984), in Proceedings *Sixth Kyoto Summer Institute on Chaos and statistical methods*, Kuramoto (Eds), 211, Springer series in synergetics
28. Joulin, G., Cambray, P., (2000) *On the scaling law for coarsening cells of premixed flames*, Comb. Sci. Tech., in press
29. Blinnikov, S.I., Sasorov, P.V., (1996), *The Landau–Darrieus instability and the fractal dimension of flame fronts*, Phys. Rev. E, **53**-5, 4827
30. Gostintsev, Y.A., Istratov, A.G., Shulenin, Yu.V., (1988), *Self-similar propagation of a free turbulent flame in a mixed gaseous mixture*, Comb. Expl. Shock Waves, **24**-5, 63
31. Joulin, G., (1994), *On the nonlinear hydrodynamic instabilities of expanding flames: intrinsic dynamics*, Phys. Rev. E, **50**, 2030
32. Joulin, G., (1989), *On the hydrodynamic stability of curved premixed flames*, J. Phys. (France), **50**, 1069
33. Zel'dovich, Y.B., Istratov, B., Kidin, N.G., Librovich, V.B., (1980), *Flame propagation in tubes : hydrodynamics and stability*, Comb. Sci. and Tech., **24**, 1
34. Pelcé, P., (1988), *Dynamics of curved fronts*, Perspectives in Physics, Academic Press, Inc, Boston
35. Clavin, P., (1988), *Theory of flames*, Disorder and Mixing, 293, Guyon et al. (Eds), NATO ASI Series, Kluwer, Dordrecht
36. Cambray, P., Joulin, G., (1994), *Length-scales of wrinkling of weakly-forced, unstable premixed flames*, Comb. Sci. and Tech., **94**, 405
37. Cambray, P., Joulain, K., Joulin, G., (1994), *Mean evolution of wrinkle wavelengths in a model of weakly-turbulent premixed flame*, Comb. Sci. and Tech., **103**, 265
38. Joulin, G., (1988), *On a model for the response of unstable premixed flames to turbulence*, Comb. Sci and Tech., **60**, 1
39. Kuppervasser, O., Olami, Z., Procaccia, I., (1996), *The geometry of developing flame fronts : analysis with pole-decomposition*, Phys. Rev. Lett., **76**, 146
40. Joulin, G., Cambray, P., Joulain, K., (1994), in Proceedings of the Zel'dovich memorial , *Combustion, Detonation, Shock waves*, Russian Section of the Combustion Institute, Frolov (Ed), vol. I, 212
41. Clanet, C., Searby, G., (1998), *First experimental study of the Darrieus-Landau instability*, Phys. Rev. Lett., **80**-17, 3867
42. Istratov, A.G., Librovich, V.B., (1969), *On the stability of gasdynamic discontinuities associated with chemical reactions : the case of spherical flames*, Acta Astronautica, **14**, 453
43. Minaev, S.S., Pirogov, E.A., Sharypov, O.V., (1996), *A nonlinear model for hydrodynamic instability of an expanding flame*, Comb. Expl. Shock Waves, **32**, 481
44. Filyand, L., Sivashinsky, G.I., Frankel, M.L., (1994), *On self-acceleration of outward propagating wrinkled flames*, Physica D, **72**, 110
45. Cambray, P., Joulain, K., Joulin, G., (1996), *Coalescence problems in the theory of expanding wrinkled premixed flames*, Comb. Sci. and Tech., **112**, 271
46. Cambray, P., Joulain, K., Joulin, G., (1996), *Dynamique, propre ou bruitée, de flammes plissées en expansion*, Unpublished DRET report 93/060, ENSMA–LCD, Poitiers University (France)

47. Groff, E.G., (1982), *The cellular nature of confined spherical propane-air flames*, Comb. Flame, **48**, 51
48. Kadowaki, S., (1995), *Numerical analysis on instability of cylindrical flames*, Comb. Sci and Tech., **107**, 181
49. Liftshitz, I.M., Slyozov, V.V., (1961), *The kinetics of precipitation from supersaturated solid solutions*, J. Phys. Chem. Solids, **19**, 35
50. D'Angelo, Y., Joulin, G., Boury, G., (2000), *On model evolution equations for the whole surface of 3–D expanding wrinkled premixed flames*, Comb. Theor. Modelling, in press
51. Palm-Leis, A., Strehlow, R.A., (1969), *On the propagation of turbulent flames*, Comb. Flame, **13**, 111
52. Kuznetsov, E.A., Minaev, S.S., (1996), *Formation and propagation of cracks on the flame surface*, Phys. Lett. A, **221**, 187
53. Truffaut, J.M., Searby, G., (1999), *Experimental study of the Darrieus-Landau instability on an inverted V flame and measurement of the Markstein number*, Comb. Sci. and Tech., **149**, 35
54. Truffaut, J.M., Searby, G., Joulin, G., (2000), *Comparison of experiments and a nonlinear model equation for spatially developing flame instability*, Submitted
55. Flyvbjerg, H., (1993), *Model for coarsening froths and foams*, Phys. Rev. E, **47**-6, 4037
56. Sivashinsky, G.I., (1983), *Instabilities, pattern formation and turbulence in flames*, Ann. Rev. Fluid Mech., **15**, 179
57. D'Angelo, Y., Joulin, G., Boury, G., (2000), *Model Evolution Equation and Numerical Solution for Wrinkled Premixed Expanding 3–D Flames*, Poster presented at the $XVII^{th}$ Sitges Euroconference Conference on Coherent Structures in Classical Systems, Sitges, Barcelona, Spain (5–9 June 2000)
58. Zel'dovich, Y., Barenblatt, G.I., Librovich, V.B., Makhviladze, G.M., (1985), *The mathematical theory of combustion and explosions*, Consultant Bureau, N.Y.
59. Josserand, C., (2000), Private communication, *Josseran@lmm.jussieu.fr*
60. Germinard, J.C., Pelcé, P., (1992), *Statistical approach for radial fingering in a Hele-shaw cell*, J. Phys. II (France), **2**, 1931

Experimental Studies of Laminar Flame Instabilities

Geoff Searby and Jean-Marie Truffaut

I.R.P.H.E., Service 252, Campus Universitaire de St. Jérôme
F-13397 Marseille Cedex 20, France

Abstract. We first briefly recall the basic mechanisms controlling the thermo-diffusive and hydrodynamic stability of planar laminar premixed flames, and give the state of the theoretical analysis. We then describe recent novel experiments to observe and measure the growth rate of cellular structures on initially planar flames. The first experiment concerns the observation of the temporal growth of wrinkling on an freely propagating planar flame. A second experiment concerns the spatio-temporal growth of structures of controlled wavelength on an anchored flame. The experimental observations are compared to theoretical dispersion relation. Finally, we compare observations of the non-linear evolution to saturation with the predictions of an extended Michelson-Sivashinsky equation.

1 Introduction

Laminar premixed flames are chemical reaction waves that propagate through a gaseous mixture. The flame front separates cold non-equilibrium reactive species from hot combustion products. In general, the planar flame sheet is unstable to perturbations and will spontaneously form cellular structures. In the following paragraphs we will briefly recall the physical mechanisms leading to flame wrinkling. The reader is referred to the cited papers and references therein for more details.

1.1 Thermo-Diffusive Instability

The propagation mechanism of the flame is essentially diffusive in nature. Heat released by the chemical reaction diffuses towards the cold gas, increasing the temperature. Since the chemical reaction rate increases exponentially with temperature, the major part of the chemical reaction and heat release takes place at the hot (burnt) side of the temperature gradient. The concentration of the reactive species is low in the reaction zone, where they are consumed. Hence the chemical reaction zone is preceded by a diffusive zone having both a temperature gradient and a concentration gradient of the reactive species. The characteristic thickness of the thermal and concentration gradients is given by $\delta = D/S_L$, where δ represents the thickness of the diffusion zone, D is the thermal or species diffusivity and S_L is the laminar flame propagation speed. For hydrocarbon air flame the value of δ is typically 50 to 100 microns.

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 159–181, 2001.
© Springer-Verlag Berlin Heidelberg 2001

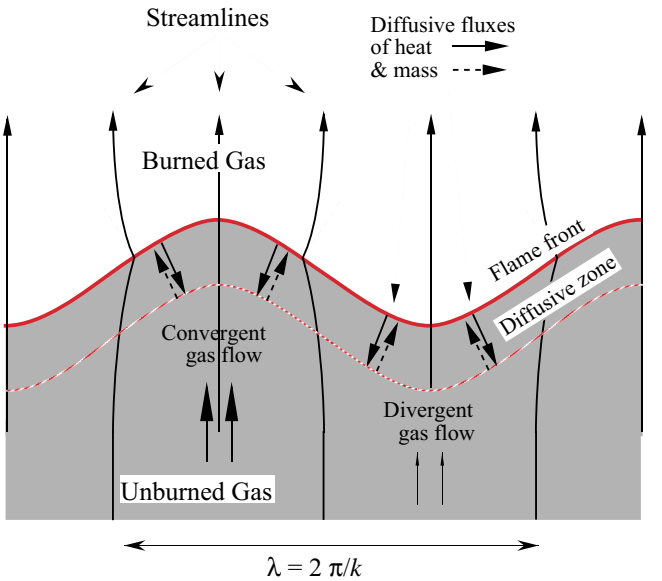

Fig. 1. Structure of wrinkled premixed flame front showing diffusive fluxes of heat and mass within the diffusive zone and hydrodynamic streamlines.

Consider a laminar flame front propagating into unburned gas whose flow velocity is equal to the flame velocity. The stability of the flame can be investigated by looking at the evolution of a harmonic perturbation of wavelength λ and wavenumber $k = 2\pi/\lambda$. If the scale of wrinkling is comparable to the flame thickness, $k\delta \approx O(1)$, the front cannot be considered as being locally plane and the temperature and species gradients become convergent or divergent at the flame front, according to the sign of the local curvature, see Fig. 1. Moreover, since the gradients of temperature and reactive species are opposed, curvature of the flame front will cause one flux to become divergent and the other to become convergent. For example a front which is convex towards the unburned gas will give rise to a divergent heat flux, which decreases the reaction rate and decreases the local flame speed. The convex front will also give rise to a convergent species flux, which increases the local reaction rate and flame speed. The net effect of the modified fluxes on the flame propagation speed depends on the ratio of the value of the thermal diffusivity to the value of the species diffusivity of the stoichiometrically deficient reactive species, i.e. on the Lewis number $Le = D_{\rm th}/D_{\rm mol}$. This is the origin of the so-called thermo-diffusive flame instability, first predicted by Zeldovich in 1944 [1,2,3]. According to this reasoning, if $Le < 1$, then the contribution of species flux dominates the change in the local burning velocity and it is easily seen that the planar flame is thermo-diffusively unstable. In the opposite case, $Le > 1$, the flame front is thermo-diffusively stable to wrinkling. Since this is a local diffusion controlled process, the growth rate (or damping rate) is proportional to the square of the wavenumber.

A different type of one-dimensional diffusive instability of premixed laminar flames (pulsating propagation), first predicted by Sivashinsky [4,5] is expected to exist for mixtures having a very high Lewis number coupled with a very high activation energy. However the theoretical threshold value of of these parameters is beyond the range of real flames and this type of instability has never been observed experimentally.

1.2 Hydrodynamic Instability

The existence of an intrinsic hydrodynamic instability of flame fronts was first recognised by Darrieus in 1938 [6] and independently by Landau in 1944 [7]. A more complete description can be found in the work of Zeldovich et al. [8]. The hydrodynamic stability of the flame is investigated by considering a harmonic perturbation whose wavelength is much greater than the flame thickness, $k\delta \ll 1$. The flame front is thus treated as a hydrodynamic discontinuity propagating towards the unburned gas with a constant normal flame speed, S_L. Thermal gas expansion through the flame increases the component of the gas velocity normal to the flame front, but the tangential component of the gas velocity is unchanged. If the stream lines are inclined, gas expansion will deviate them towards the downstream normal to the flame, see Fig.1. For usual flames, the Mach number w.r.t. the flame speed is small, so the flow is quasi-incompressible and the pressure gradients and vorticity associated with the deviation of the downstream streamlines will also deviate the upstream flow. This is thus a non-local effect. In the case of a wrinkled flame, this self induced modification of the flow causes the flow to decelerate upstream of a region which is convex towards the unburned gas and to accelerate in front of a region which is concave. In this approximation, $k\delta \ll 1$, a wrinkled flame is hydrodynamically unstable at all wavelengths.

The growth rate, σ, of the instability can be written in a non-dimensional form as :

$$\sigma\tau_t = k\delta\,\Omega(E), \tag{1}$$

where

$$\Omega(E) = \frac{E}{E+1}\left(\sqrt{\frac{E^2+E-1}{E}} - 1\right), \tag{2}$$

k is the wave number of the wrinkling, $\tau_t \equiv \delta/S_L$ is the flame transit time, E is the ratio unburned to burned gas densities, $E = \rho_{\mathrm{u}}/\rho_{\mathrm{b}}$ and $\Omega(E)$ is a positive function of order unity, vanishing for $E = 1$.

1.3 Complete Description of Laminar Flame Stability

The effects of diffusion and of gas dynamics are not exclusive and several authors have given a combined description of the two phenomena [9,10,11,12,13]. In the latter references, the approximation of high activation energy ($\beta \approx 10$), see (7), and multi-scale asymptotics are used to solve the problem at three physically

distinct scales, which are the inner scale of the chemical reaction zone (scale δ/β), the scale of the diffusive zone (scale δ) and the outer scale of the hydrodynamic zone (scale k^{-1}). The effect of gravity acting on two fluids of different density, burned and unburned gas, is also included on the analysis. The effect of gravity can help stabilise flames when the light burned gas is above the heavy unburned gas. The most complete description of laminar flame stability is given in a little known paper of Clavin and Garcia [14]. These authors solve the dynamics of small amplitude wrinkling of premixed flames including the effect of temperature dependent diffusion coefficients. They find that the dispersion relation for the growth rate, σ, of small amplitude wrinkling with wave number k is given by :

$$(\sigma\tau_{\rm t})^2 A(k) + \sigma\tau_{\rm t} B(k) + C(k) = 0, \tag{3}$$

The growth rate of the instability is given by the real part of σ. The wavenumber dependent coefficients in (3) are given by :

$$\begin{aligned}
A(k) &= \frac{E+1}{E} + \frac{E-1}{E} k\delta \left(Ma - J\frac{E}{E-1} \right), \\
B(k) &= 2k\delta + 2E\,(k\delta)^2\,(Ma - J)\,, \\
C(k) &= \frac{E-1}{E}\frac{k\delta}{Fr} - (E-1)\,(k\delta)^2 \left[1 + \frac{1}{E\,Fr}\,\left(Ma - J\frac{E}{E-1} \right) \right] \\
&\quad + (E-1)\,(k\delta)^3 \left[h_{\rm b} + \left(\frac{3E-1}{E-1} \right) Ma - \frac{E}{E-1} 2J + (2Pr-1)\,H \right].
\end{aligned} \tag{4}$$

Here, $\delta = D_{\rm th}/S_L$ is the thickness of the thermal diffusion zone, $Fr = S_L^2/(g\delta)$ is the Froude number of the flame, g is the acceleration of gravity (positive when the flame propagates downwards) and Pr is the Prandtl number. The quantities J and H are given by the integrals :

$$H = \int_0^1 (h_{\rm b} - h(\theta))\,\Delta\theta, \qquad J = (E-1)\int_0^1 \frac{h(\theta)}{1+(E-1)\,\theta}\Delta\theta, \tag{5}$$

where $\theta = (T - T_0)/(T_{\rm b} - T_0)$ is the reduced temperature, $h(\theta)$ is the thermal diffusivity times density ($\rho D_{\rm th}$) at temperature θ, normalised by its value in the unburned gas, and $h_{\rm b}$ is the value of $h(\theta)$ in the burned gas. The integrals J and H express the effect of the temperature dependence of the diffusivity. The notation used here is slightly different to that in the paper of Clavin and Garcia. In particular the quantity $l\mathcal{D}$ in the Clavin–Garcia paper is re-written, using their equation $(24')$, as $l\mathcal{D} = 2\,[(E-1)Ma - J\,E]$. The above relations were obtained in the linearised limit of long wavelength low frequency perturbations $k\delta \ll 1$, $\sigma\tau_t \ll 1$ with small amplitude of wrinkling. The expressions are accurate to order $(k\delta)^2$ and $(\sigma\tau_t)^2$.

The sensitivity of the local burning velocity to curvature and stretch is contained in the Markstein number, Ma, [9]. When the Markstein number is positive, which is generally the case, the effect of curvature is to decrease the burning velocity of a region which is convex towards the unburned gas, implying that

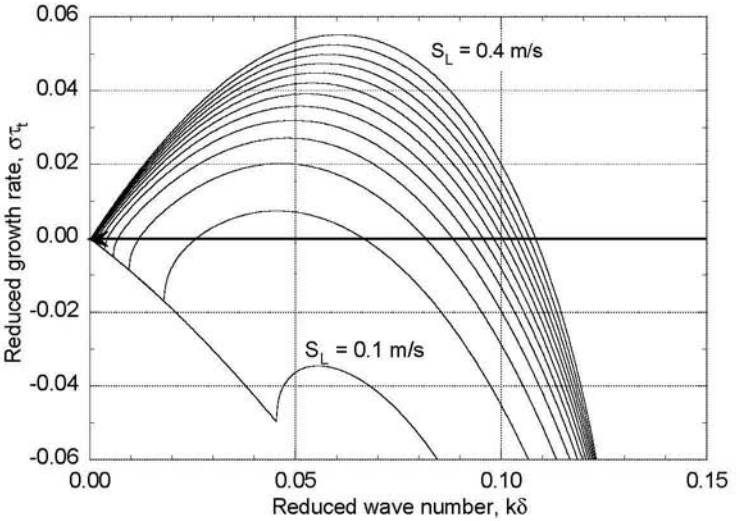

Fig. 2. Reduced growth rate of Darrieus–Landau instability plotted as a function of reduced wavenumber for flame speeds in the range $0.1 \leq S_L \leq 0.4\,\mathrm{m/s}$. (lower to upper in steps of 0.025 m/s). Other parameters appropriate for lean propane flames. The data needed in equations (4) and (5) were taken from the CHEMKIN package [15]. Ma= 4.5 [16].

the flame is thermo-diffusively stable. Clavin and Garcia [14] have given an expression for the Markstein number for the simplified case of an overall one-step chemical reaction controlled by an Arrhenius law :

$$Ma = \frac{E}{E-1}J + \frac{\beta}{2}(Le-1)\int_0^1 \frac{h(\theta)\,\ln(1/\theta)}{1+(E-1)\theta}\Delta\theta. \tag{6}$$

In (6), β is the reduced activation energy, also known as the Zeldovich number :

$$\beta = \frac{E^*}{\mathrm{k}T_\mathrm{b}}\frac{T_\mathrm{b}-T_0}{T_\mathrm{b}}, \tag{7}$$

where E^* is the activation energy of the reaction and k is Boltzmann's constant. The first term on the r.h.s. of (6) arises from transverse *convective* transport of heat and species within the diffusive thickness of the wrinkled flame. The second term arises from transverse *diffusive* transport of heat and mass. Contrary to the reasoning presented in section 1.1, (6) shows that, when convection within the flame thickness is also considered, a flame can be thermo-diffusively stable ($Ma > 1$) even for Lewis numbers smaller than unity.

The approximation of one-step chemistry is not good and in general (6) should be considered only as a rough approximation. It is generally necessary to obtain Markstein numbers from experimental measurements or from direct numerical simulations. Typical plots of the real part of the reduced growth rate, $\sigma\tau_\mathrm{t}$, calculated for propane air flames are given in Fig. 2. For small wave numbers, the growth rate first increases linearly with the wave number, as predicted

by Darrieus and Landau. Then, at larger wave numbers, the growth rate decreases with the square of the wavenumber, due to thermo-diffusive effects. The effect of gravity is to shift the growth rate curve downwards and to reduce the range of instable wavenumbers. The effect is only significant for slow flames, $S_L < 20\,\mathrm{cm/s}$. For sufficient slow flames, $S_L < 12\,\mathrm{cm/s}$, the effect of gravity can stabilise downwards propagating flames at all wave numbers [16,12].

This linear analysis is expected to be correct during the linear part of the growth of the Darrieus–Landau instability. However, until recently, there has been no direct experimental verification of the theory. The reason for this lies in the difficulty of producing the initial condition of a planar, freely propagating, unstable flame. The characteristic growth time of the instability is typically 20–50 ms, which is short compared to the time needed to establish a freely propagating flame of finite dimensions. It follows that the Darrieus–Landau instability has nearly always been observed in the phase of non-linear saturation.

2 Experimental Verification on Planar Flames

The first direct experimental verification of (3) was performed in 1998 by Clanet and Searby [17]. They used a novel technique of acoustic restabilisation to produce a perfectly planar laminar flame which is otherwise intrinsically unstable. The imposed acoustic field was then removed on a short time scale, (about 1 ms), and the unconstrained growth of the Darrieus–Landau instability was observed. The contents of this section will describe their work.

2.1 Effect of Acoustic Field on a Laminar Flame

The effect of an imposed acoustic field on a freely propagating premixed laminar flame front has been given by Searby & Rochwerger [18]. In the approximation $a\,k \ll 1$, where a is the amplitude of the wrinkling, these authors have shown that the dynamics of a flame front in an imposed acoustic field, characterised by the frequency, ω_a, and displacement velocity, u_a, is dominated by the effect of the periodic acoustic acceleration acting on the flame seen as a thin interface between two fluids of different density. The dynamics of the system can be obtained from (3) and (4) by introducing this periodic acceleration into the Froude number, along with the acceleration of gravity :

$$Fr^{-1} = \frac{g\delta}{S_L^2} - \omega_a u_a \cos(\omega_a t). \tag{8}$$

Writing $\alpha(y,t) = a\exp(\sigma t - \mathrm{i}ky)$, where α is the amplitude of wrinkling at wave number k, it can be seen that (3) is the time Fourier transform of the following equation of motion for α :

$$A(k)\,\frac{\partial^2\alpha}{\partial t^2} + B(k)\frac{\partial\alpha}{\partial t} + \left[C(k) - C_1(k)\cos(\omega_a\ t)\right]\alpha = 0. \tag{9}$$

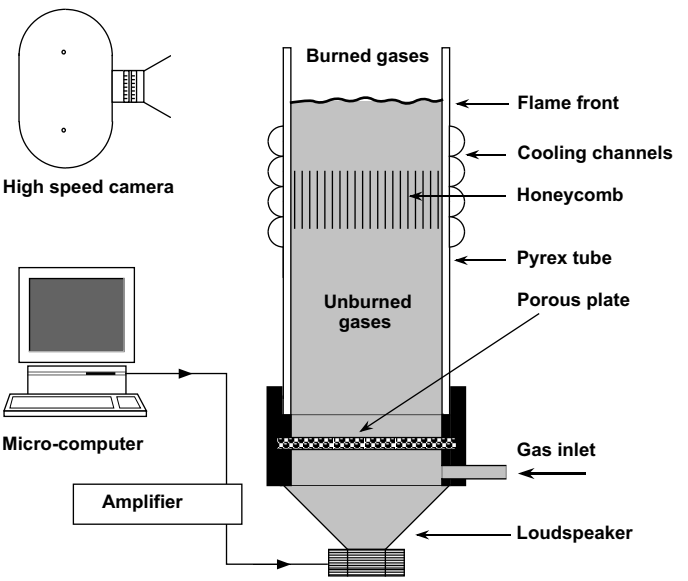

Fig. 3. Experimental set up used to prepare an acoustically stabilised planar flame.

The wavenumber dependent coefficients $A(k)$, $B(k)$ and $C(k)$ are given in (4). The acoustic acceleration has introduced an extra term $C_1(k)$:

$$C_1(k) = \frac{E-1}{E} k\delta\, \omega_a u_a \left[1 - k\delta \left(Ma - J\frac{E}{E-1}\right)\right] \tag{10}$$

Equation (9) is formally equivalent to a parametrically driven damped harmonic oscillator in which the acoustic acceleration appears as the driving force. This equation, (9), can be reduced to the Mathieu equation [19] by a simple substitution of variables [9]. The solutions of this latter are known to present several regions of instability separated by a stable domain. Details are given in [18]. It may be noted that the flame front does not have a single characteristic frequency, but a continuum associated with the continuum of wavelengths that can be excited on its surface. At excitation frequencies of a few hundred Hertz and for moderate flames speeds less than 20 cm/s, the flame is found to be stable at all wavelengths for a narrow range of acoustic excitations having acoustic displacement velocities $u_a \approx 4S_L$. This corresponds to an acoustic intensity of about 140 dB. This acoustic stabilisation window has been used to maintain an intrinsically unstable flame in a planar state, prior to observing the growth of the Darrieus–Landau instability.

2.2 Experimental Set-Up for Acoustically Stabilised Flame

The experimental apparatus is shown in 3. The premixed gases are fed into the bottom of a Pyrex glass tube 100 mm diameter and 400 mm long, just below a

50 µm porous plate whose role was to laminarise the flow. The flame was held stationary in the laboratory frame, about 50 mm below the tube exit, by careful adjustment of the gas flow rate. A 2 mm aluminium honeycomb structure 40 mm long was placed a few centimetres upstream from the flame to help maintain a laminar and homogeneous gas flow. A helical cooling tube was wound round the outside of the Pyrex tube, below the flame front, in order to prevent wall heating by heat conduction and radiation from the flame. The Pyrex tube was closed at its lower extremity by a loudspeaker which was fed from a micro-computer used as a programmable signal generator. The acoustic impedance of the porous plate is high and the Pyrex tube behaved as an open–closed resonator. It was excited in the 1/4 wavelength longitudinal mode (230 Hz). The natural acoustic damping time (≈12 ms) was not always short compared to the growth time of the Darrieus–Landau instability. To overcome this limitation, at the start of a measurement, the damping of the acoustic standing wave was increased artificially by injecting one cycle of a signal in phase opposition with the pressure in the tube. The luminous emission from the flame was recorded by an intensified high speed cine camera.

Experiments were performed with lean propane-air mixtures with equivalence ratios in the range $0.56 \leq \phi \leq 0.67$, corresponding to burning velocities in the range $11.5\,\mathrm{cm/s} \leq S_L \leq 20\,\mathrm{cm/s}$. For leaner flames, the front was intrinsically stable at all wavelengths. For faster flames, a flat laminar flame could not be obtained by this method. In order to control the wavelength and orientation of the cellular structures that developed when the acoustic stabilisation was removed, the upstream gas flow was perturbed by placing an array of parallel wires, 2 mm in diameter, on the downstream face of the honeycomb. The object of this scheme was to excite purely 2-D cells at a chosen wavelengths of 2 cm, close to the most unstable wavelength, see Fig.2. The luminous emission from the flame front was filmed edge-on in a direction parallel to the axis of the cells. Fig.4 shows a sequence of typical images taken from a high speed film during the growth of the Darrieus–Landau instability. The time $t = 0$ corresponds to the instant at which the acoustic field was removed.

The apparent thickening of the flame, particularly at high cell aspect ratios indicates the presence of slight three-dimensionality of the wrinkling. The peak to peak amplitude of the wrinkling was measured on digitized images and plotted in semi-log coordinates as shown in Fig.5.

The large scatter of the points in the early stages of the growth arise from the small amplitude of the cells, of the order of the apparent flame thickness. The non linearity at long times indicates the onset of saturation of the instability. This nonlinearity of the shape of the cells is clearly visible in Fig.4 after 140 ms. These points were systematically eliminated before data reduction to obtain the growth rate. The points were fitted to an exponential function of the form :

$$a(t) = \frac{1}{2}\left[\left(a_o + \frac{v_0}{\sigma}\right)e^{\sigma t} + \left(a_o - \frac{v_0}{\sigma}\right)e^{-\sigma t}\right], \tag{11}$$

which is the general solution of $\partial^2 a/\partial t = \sigma^2 a$ with the initial conditions $a(0) = a_0$ and $\partial a(0)/\partial t = v_0$. Here, v_0 is the rate of increase of the wrinkling at time $t =$

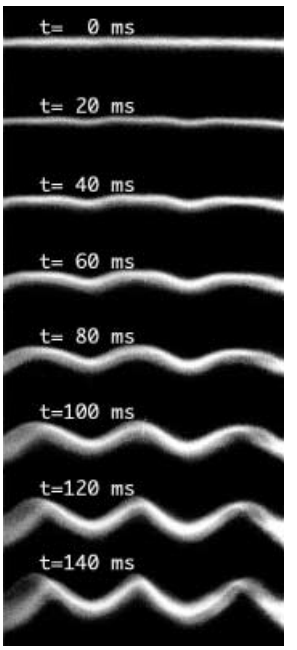

Fig. 4. Images taken from a high speed film of growth of instability. Framing rate 500 i/s. Wavelength 2 cm. Flame speed 11.5 cm/s.

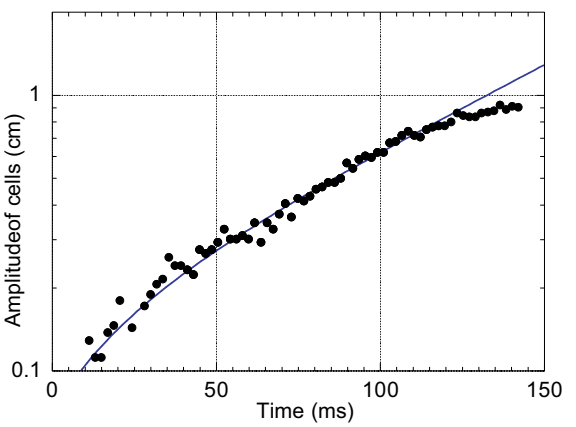

Fig. 5. Semi-log plot of the amplitude of the cellular structures of Fig.4. Fitted curve from (11).

0, supposed equal to the measured peak to peak velocity modulation produced by the wires in the flow. The precision of the growth rate obtained by this procedure was $\pm 5\,\mathrm{s}^{-1}$. These experimentally measured growth rates are plotted as a function of laminar flame speed in Fig.6.

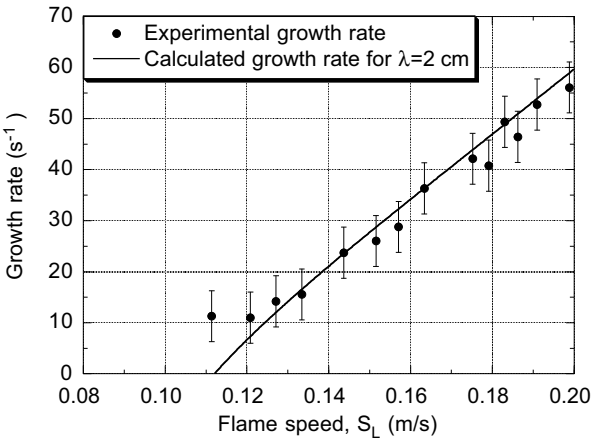

Fig. 6. Comparison of experimental and calculated growth rates. The flame speeds, S_L, were taken from [20] and gas expansion ratios were calculated using the CHEMKIN package [15].

All measurements were made at a fixed (forced) wavelength of 2 cm. The full line in Fig.6 shows the calculated growth rates obtained from the dispersion relation (3). A Markstein number of 4 was found to give best agreement with the experimental data. This value of 4, for lean propane–air mixtures may be compared with the value 5 found by Tseng *et al.* [21] and the value 4.3 found previously by Searby and Quinard [16].

3 Experimental Verification of Growth Rate on Inclined Flames

In the previous section we described an experimental technique that has been used to observe the growth rate of the Darrieus–Landau instability on an initially planar flame. Because of limitations inherent to the technique, the observations were limited to relatively slow flames, $S_L < 20\,\mathrm{cm/s}$ and to observation at a wavelength close to that of maximum growth rate. In this section we will describe a different experimental technique used recently by Truffaut and Searby [22] to measure the growth rate of the instability on an inclined flame over a wider range of wavenumbers and flame speeds. In these experiments a small amplitude oscillation, at frequency f, was imposed at base of the inclined flame and the resulting structure was convected downstream by the tangential component of the gas flow. The wavelength of the structure is $\lambda = u_\mathrm{t}/f$ where u_t is the tangential convection velocity. In this configuration the growth of the instability is spatio-temporal and not simply temporal. However, the spatial growth rates were converted to temporal growth rates using a Lagrangian time obtained from the the displacement velocity of the wrinkles, u_t. This conversion is expected to be valid if adjacent cells have (nearly) the same size. This condition is satisfied

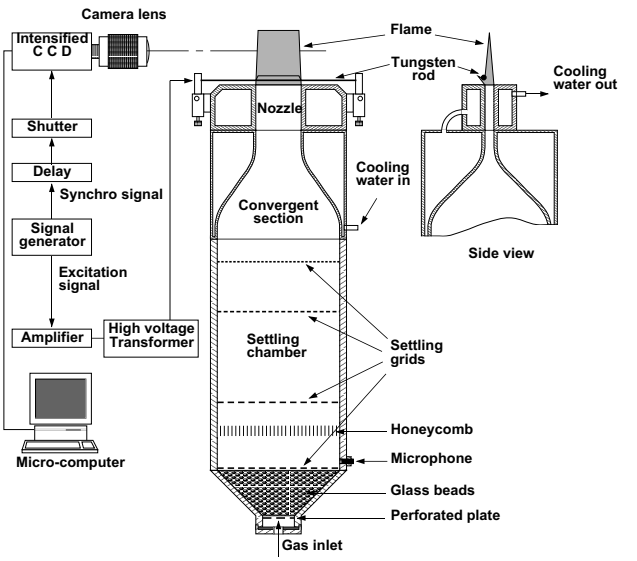

Fig. 7. Schematic diagram of the inverted-'V' burner and excitation system.

when the growth time is long compared to the time taken to convect the pattern a distance of one wavelength downstream, i.e. for $\Re[\sigma]^{-1} \gg f^{-1}$. Thus (3)–(5) were again used to analyse the results of the experimental measurements.

Equations (3)–(5) are strictly valid only for planar flames propagating upwards or downwards. The case of an inclined flame in a gravity field has been treated theoretically by Garcia and Borghi [23]. The effect of gravity is to introduce a bulk force with a component parallel to the flame front and the problem becomes asymmetric. In a simplified analysis, Garcia and Borghi find that 2-D wrinkles, whose axes of wrinkling are horizontal, can can have a finite propagation velocity, different from the convective velocity. According to their analysis, this propagation velocity goes to zero for horizontal and for vertical flame fronts. In the experiments of Searby and Truffaut, the flame fronts are close to vertical ($< 5°$) and moreover the Froude number of the flames, $Fr = S_L^2/(g\delta)$ is always very large (> 90). The effect of gravity is then expected to be negligible. Experimentally, it was found that the measured displacement velocity of the wrinkles was equal to the velocity of the tangential gas flow.

3.1 Inverted-'V'Burner and Flame Excitation Device

A laminar slot burner was used to produce a two-dimensional inverted-'V' premixed flame and an the electrostatic deflection system was used to impose the wavelength of the perturbation.

The experimental apparatus is shown schematically in Fig.7. The premixed propane-air-oxygen gas was fed to the bottom of a burner designed to produce a laminar 'top-hat' velocity profile at the exit. The burner comprised a divergent section, a settling chamber, a convergent section and a nozzle. The 60 mm

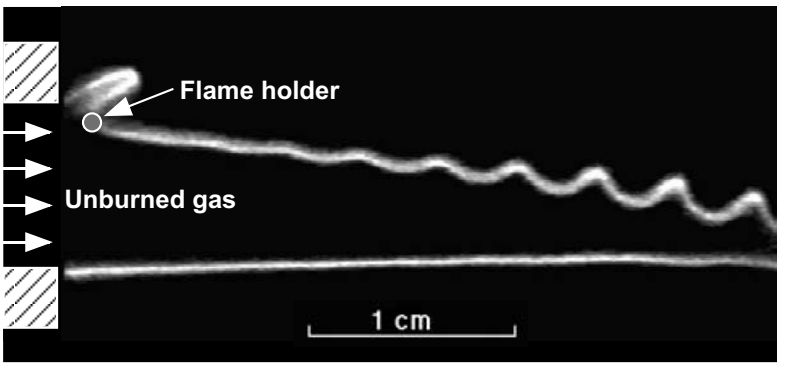

Fig. 8. Instantaneous image of growth of Darrieus–Landau instability of a propane–air–oxygen flame, equivalence ratio = 1.33, 28% oxygen, flow velocity 8.56 m/s, excitation frequency 2100 Hz.

high divergent section was filled with 4 mm diameter glass beads to break up the incoming flow. The flow was then laminarised in a 140 mm square settling chamber containing an aluminium honeycomb, followed by three metal grids of decreasing mesh size. Finally a 2-D convergent section, with a contraction ratio of 30:1, accelerated the flow up to the 8*80 mm exit section and reduced the residual velocity fluctuations to a small fraction of the mean flow velocity. The 1:10 aspect ratio of the exit provided a 2-D inverted-'V' flame. The flame was perturbed by end effects for less that 10 mm at each extremity. The thickness of the viscous boundary layers was less than 1.5 mm and the transverse velocity profiles were flat to better than 1% over the central region of 5 mm. The residual turbulence was less than 0.5%. In order to minimise the development of a thermal boundary layer near the walls, both the convergent section and nozzle were water-cooled at the temperature of the unburnt mixture (20°C).

One side of the flame was anchored on a thin tungsten rod, 0.6 mm diameter, placed just above and parallel to the burner exit, see Fig.8. The laminar flame was perturbed on one face by applying an alternating high voltage between the rod and the burner exit. The resulting electric field displaces the flame locally and produced a periodic sinusoidal 2-D wrinkle on the flame front as shown in Fig.8. The axis of the wrinkle was parallel to the burner slot. The wrinkle was convected downstream by the gas flow and its amplitude was observed to increase exponentially. This system permitted a precise control of both the initial amplitude and of the wavelength of the wrinkle through control of the amplitude and of the frequency of the applied signal. Typical ranges of voltage and frequency were 2–4 kV and 1–4 kHz respectively.

A similar electrostatic technique was first used by Polanyi and Markstein in 1947 [24]. However the present system is slightly different, in that Markstein placed an deflection electrode in the burned gas, whereas we used the electrode as a flame holder. We have observed that, for a given voltage, the amplitude of small wavelength wrinkles was much greater when the flame was excited via a flame-

holder, probably because the electric field acts more locally on the flame front. The deflection of a flame in an electric field is attributed to a body force produced by momentum transfer between charged particles and neutral molecules. Some authors [25] and [26] have argued that the momentum transfer from heavy ions dominates momentum transfer from electrons, giving rise to a non-zero body force in the direction of the electric field.

Experiments were performed on propane–air flames and propane-air-oxygen flames with equivalence ratios in the range $1.05 \leq \phi \leq 1.33$. The corresponding laminar flame velocities were in the range $0.43\,\mathrm{m/s} \geq S_L \geq 0.27\,\mathrm{m/s}$ [20] for the propane-air flames. The velocities were 0.69 m/s and 0.51 m/s for flames with 28% oxygen at equivalence ratios of 1.05 and 1.33. The gas flow velocities were 7.4 m/s, 6.05 m/s and 8.56 m/s.

3.2 Data Acquisition

The wrinkled flame front was observed using a short exposure intensified CCD camera, viewing parallel to the the long edge of the burner slot. In this configuration the flame is a long 3-D object. Since we were interested in imaging a cross section of the flame, we used an optical system having a short depth of focus and a high magnification ratio. The camera was focussed in the centre of the burner slot and the depth of focus was roughly 2 mm. This value also represented the uncertainty in the knowledge of the position of the object plane. The resulting relative uncertainty in the magnification of the image is less than 1%. An inconvenient of this optical system is some unsharpness of the image due to the superposition of the out-of-focus contours from other distances. However the resolution of the images obtained by this method was acceptable, see Fig.8.

The video camera was triggered at a frequency close to 50 Hz by a pulse generator synchronized to an appropriate sub-harmonic of the excitation signal. An adjustable delay allowed us to take images at different phases of the excitation. For all images, the gate-time of the intensifier was less than 100 µs and the optical gain was adjusted manually so as to obtain non-saturated images. The images were digitized at a resolution of 760*570 pixels. The spatial calibration, obtained by imaging a grid placed in the object plane, was 52.6 µm per pixel.

In the experiment described here, the flame front was not freely propagating, but attached to a flame holder in the proximity of a companion front, forming an inverted-'V' flame. In this situation, care must be taken to ensure that the observed perturbations were growing freely. Firstly it was ensured that residual and external flow perturbations were so small that only the imposed wavelength appears on the flame front. This was verified by noticing that, in the absence of electric excitation there was no detectable motion of the flame front. The residual perturbations were so small that they did not have time to grow to a measurable amplitude before they were convected to the flame tip. Secondly it was ensured that the two flame fronts did not interact. It is known that the hydrodynamic field ahead of a wrinkled premixed flame is modified over a characteristic distance equal to k^{-1} [13]. If the distance between the two fronts is smaller than this value, the instability developing on one side of the flame

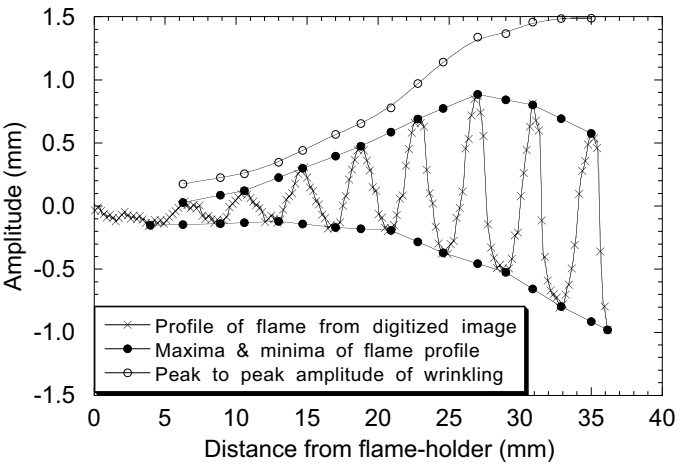

Fig. 9. Plot of flame profile from Fig.8 showing method of measuring amplitude of wrinkling and corresponding peak to peak amplitude.

can be influenced by the presence of the other flame. To avoid this problem the angle between the two sides of the flame was kept small and their spacing, in the field of visualization, was kept large compared to k^{-1}. For this reason the measurements were also restricted to wavelengths less than 6 mm. The absence of interaction between the two flame fronts can be checked from images such as Fig.8. The lower side of the flame shows no perturbation from the wrinkled side of the flame, except at the far right side of the image, close to the flame tip, where the amplitude of the wrinkling is already saturated. It may concluded that the contrary is also true, the wrinkled side of the flame was not affected by the presence of the unperturbed side.

3.3 Data Processing

Images of the flame, such the one shown in Fig.8, were processed on a desktop computer to obtain the amplitude and the wavelength of the wrinkles as a function of the distance downstream. A program searched for the brightest pixel on each vertical line of pixels in a sub-window containing the excited side of the flame. The program then fitted a parabola through the intensity at this point and at two other points on each side of the maximum. The flame position was given by the position of the maximum of the fitted curve. This algorithm gave the position of the flame front as a function of the distance from the flame holder. A typical plot is shown in Fig.9. After low pass filtering of the curve, for each maximum (and minimum), the value corresponding to excitation in phase opposition was approximated by interpolation of the preceding and following minima (or maxima), as shown in Fig.9. In this way, the peak to peak amplitude and the wavelength of the wrinkle was obtained for each extremum position. For each frequency of excitation this operation was repeated for 4 different phases of the

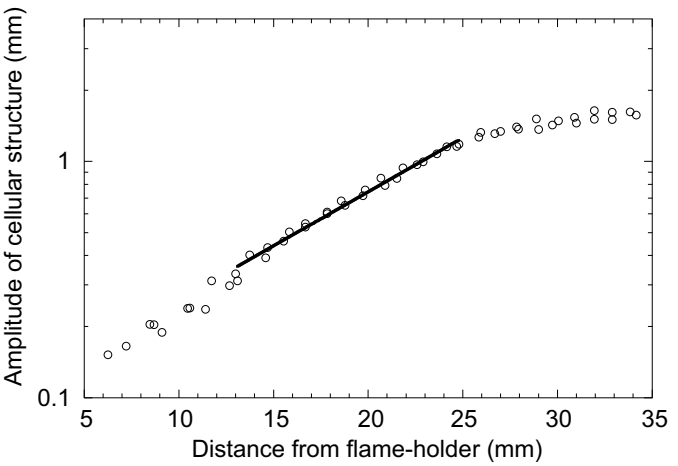

Fig. 10. Semi-log plot of peak to peak amplitude of flame wrinkling obtained from Fig.9.

signal. The peak to peak amplitude was then plotted in semi–log coordinates, as shown in Fig.10.

The spatial growth rate of the instability σ_x was obtained by fitting this curve with an exponential function. The initial points, close to the flame holder, were influenced by the presence of the rod and also by the presence of the electric field. The amplitude of the wrinkles in this region was of the order of the apparent flame thickness $\delta \approx 0.2\,\mathrm{mm}$. These points were systematically ignored, along with the downstream points in the region of amplitude saturation where the wrinkles are cusped. The amplitude of the wrinkling was measured with a precision of the order of $\delta \approx 0.2\,\mathrm{mm}$, so there was a resulting uncertainty in the values of the fitted parameters. The spatial growth rate was thus measured with a precision which ranged from 10% to 30%, depending on the size of the wrinkles.

3.4 Results

The experimental results showed that the wrinkles produced at the base of the flame were convected downstream at a constant velocity, u_t, equal to the tangential flow velocity to within experimental error. The wavelength of excitation is thus given by $\lambda = u_t/f$, where f is the excitation frequency. We have not observed any tendency for the wrinkles to have a finite propagation velocity with respect to the flow, as suggested in the work of Garcia and Borghi [23]. The amplitude of the wrinkles was observed to increase exponentially up to the onset of saturation, visible in Figs 8, 9 and 10.

The spatial growth rate was converted to a temporal growth rate from knowledge of the convection velocity of the structures, $\sigma = \sigma_x u_t$. The non-dimensional temporal growth rates were plotted as a function of the non-dimensional wave-

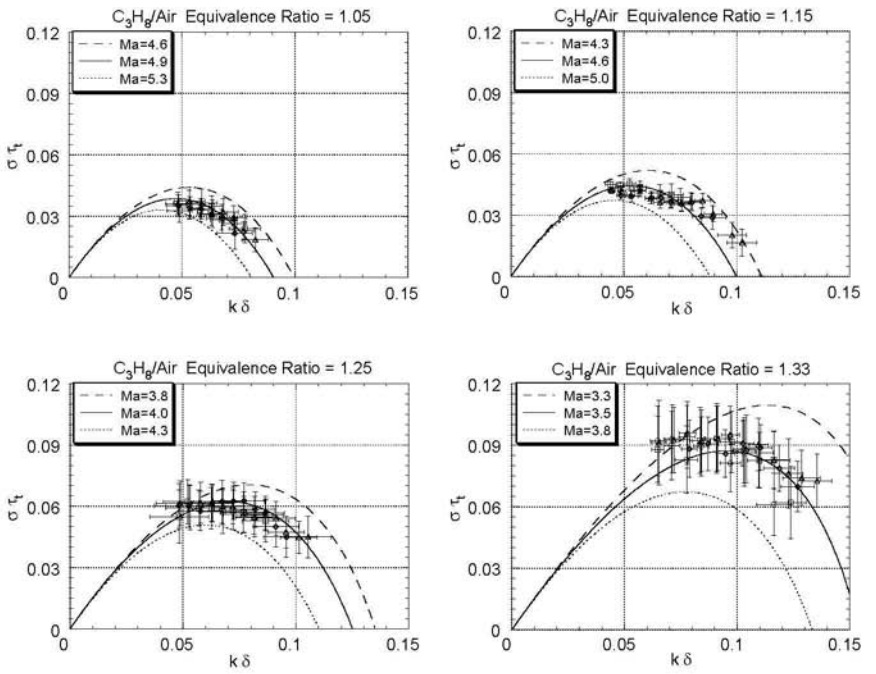

Fig. 11. Measured growth rate of Darrieus–Landau instability as a function of imposed wavenumber for propane–air flames at four different equivalence ratios. The curves are calculated from (3).

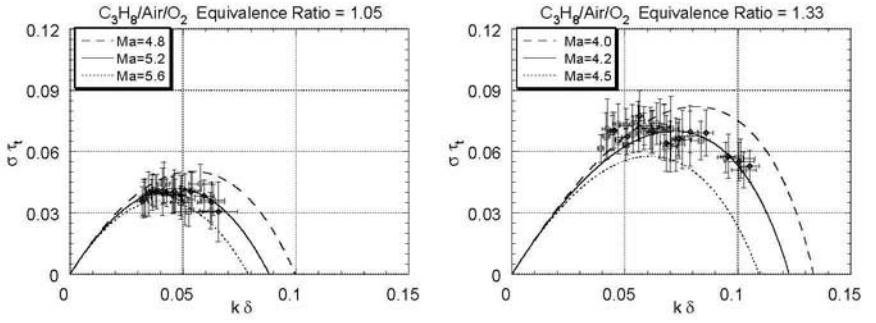

Fig. 12. Measured growth rate of Darrieus–Landau instability as a function of imposed wavenumber for propane flames with 28% oxygen at two different equivalence ratios. The curves are calculated from (3).

number in Fig. 11. The dimensional growth rate was typically 150–300 s^{-1} for propane air flames. The minimum value of f/σ was 3. Similar plots for the oxygen enriched flames are shown in Fig.12.

These experimental results are compared to the theoretical dispersion relation (3) of Clavin and Garcia [14] where the acceleration of gravity been put equal to zero. This approximation is justified because the Froude number, based

on the flame speed and gravity, is always greater than 90 in these experiments. The numerical data used for the calculations is listed in table 1. The Markstein

Table 1. Numerical values used to calculate the curves in Figs.11 and 12

	21% O_2				28% O_2	
Φ	1.05	1.15	1.25	1.33	1.05	1.33
S_L m/s	0.43	0.41	0.35	0.27	0.69	0.51
T_b K	2067	2082	2066	2033	2334	2326
δµm	47.6	49.6	57.7	74.4	28.9	38.1
τ_tµs	111	121	165	276	41.8	74.8
E	7.05	7.11	7.05	6.94	7.97	7.94
J	3.97	3.98	3.94	3.87	4.36	4.29
H	1.108	1.106	1.089	1.067	1.198	1.164

number of each flame has been treated as an unknown parameter. For each flame we have plotted the theoretical curve for the Markstein number that agrees best with the measured results. To give an idea of the variation of the theoretical curves with this parameter, we have also plotted curves for the best values of $Ma \pm 0.3$ or ± 0.2 according to sensitivity. It can be seen that the theoretical curve and the experimental measurements agree to within experimental uncertainty. It can also be seen that, for large wavenumbers, the measured growth rate of the instability decreases with increasing wavenumber, as predicted by theory including the effects of both hydrodynamics and diffusion. These experimental values of the Markstein number are comparable with other values given in the literature for rich propane–air flames [27,21].

These experiments were limited to wave lengths less than 6 mm to avoid interaction between the two flame fronts. This constraint imposed a lower limit on the reduced wavenumber, $k\delta$, which was thus limited to the range 0.04–0.15. This value is smaller than unity, but not very much smaller, so the results are at the limit of the domain of validity of linearized asymptotic flame theory ($k\delta \ll 1$). It may also be noted that the maximum reduced growth rate, $\sigma\tau_t$, is 0.1 for the richest propane–air flame. The results are thus also at the limit of validity low of the low frequency assumption.

4 Nonlinear Saturation of the Instability

In the previous sections we have been concerned with the small amplitude linear growth rate of the instability. It was seen that the growth rate is well described by linearised asymptotic laminar flame theory. In this section we will be concerned with the non linear effects leading to the saturation of the amplitude of the structures, see figs 8 and 10.

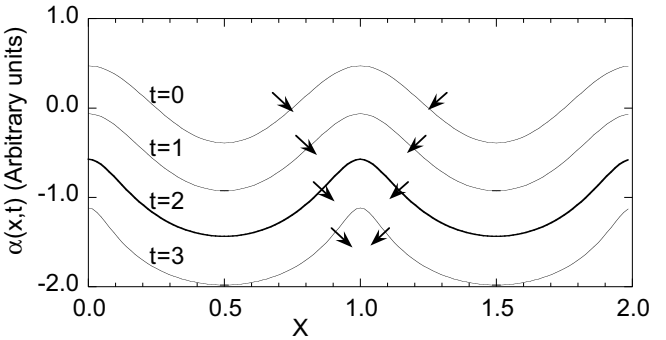

Fig. 13. Propagation of wrinkled flame front showing cusp formation.

The origin of the saturation mechanism is easily understood qualitatively. Consider a wrinkled flame, flat on average, propagating into a quiescent fresh mixture, as shown in Fig.13. At all points, the flame propagates in the direction of the local normal to the flame front. When the aspect ratio of the structures formed by the instability is not small, the local direction of propagation of the flame front is significantly different from the average plane of the flame. The evolution of the flame shape can be obtained geometrically using Huygens construction. It is easy to see that an initially sinusoidal structure will be distorted. The radius of curvature of zones which are convex towards the unburned gas will increase, whereas the radius of curvature of concave zones will decrease until a cusp is formed. Sivashinsky [10,28] has derived a nonlinear differential equation to describe the evolution of the the shape of the flame front. This equation, now known as the Michelson-Sivashinsky equation, can be written in a simplified dimensional form :

$$\frac{\partial \alpha}{\partial t} = \frac{(E-1)}{E}\frac{S_L}{2}\left(\frac{1}{k_n}\frac{\partial^2 \alpha}{\partial x^2} + I(\alpha)\right) - \frac{S_L}{2}\left(\frac{\partial \alpha}{\partial x}\right)^2, \tag{12}$$

$$I(\alpha) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \alpha(x,t)\exp\left(\mathrm{i}k(x-\xi)\right)\mathrm{d}\xi\,|k|\mathrm{d}k.$$

In (12), $\alpha(x,t)$ is the local position of the flame front with respect to some reference plane, x is the transverse coordinate and k_n is the neutral wavenumber for which $\sigma(k_n) = 0$. The linear operator $I(\alpha)$ multiplies α by $|k|$ in wavenumber space. It corresponds to the Darrieus–Landau instability. The second derivative, $\partial^2\alpha/\partial x^2$, gives the change in local burning velocity with curvature. The non linear term, $(\partial\alpha/\partial x)^2$, arises from inclination of the flame front and together with the curvature term, $\partial^2\alpha/\partial x^2$, will lead to saturation. Equation (12) was obtained in the limit of Lewis number close to unity, $Le - 1 \ll 1$, and small gas expansion ratio, $E \ll 1$. The acceleration of gravity is not considered here. In real flames the gas expansion ratio is not a small parameter, typically $E \approx 6$. Joulin and Cambray [29] have proposed the following phenomenological extension to deal with deeply wrinkled flames when the gas expansion is not small compared

to unity :

$$\frac{\partial \alpha}{\partial t} = S_L\, \Omega \left(\frac{1}{k_n} \frac{\partial^2 \alpha}{\partial x^2} + I(\alpha) \right) - \frac{S_L}{2} a \left(\frac{\partial \alpha}{\partial x} \right)^2, \tag{13}$$

where $\Omega(E)$ is the constant of proportionality found by Landau, given by (2) in section 1.2 and a is a phenomenological constant,

$$a(E) = \frac{2E}{E+1}. \tag{14}$$

Equation (13) reduces to (12) in the limit $E - 1 \ll 1$.

It is known that equation (13) (or (12)) has an exact solution for $\alpha(x,t)$ $2\pi/k$ periodic [29,30]

$$\alpha(x,t) = -A(t) - \frac{2\Omega}{a\, k_n} \log \left[1 - \frac{\cos(kx)}{\cosh\{k\mathrm{B}(t)\}} \right]. \tag{15}$$

The functions $\mathrm{A}(t)$ and $\mathrm{B}(t)$ are the solutions to the coupled ordinary differential equations [29] :

$$\frac{\mathrm{dA}}{\mathrm{d}t} = S_L \left[\left(\frac{2\Omega}{a} \frac{k}{k_n} \right)^2 \frac{\exp(-2k\mathrm{B})}{1 - \exp(-2k\mathrm{B})} \right], \tag{16}$$

$$\frac{\mathrm{dB}}{\mathrm{d}t} = S_L \Omega \left[\frac{k}{k_n} \coth(k\mathrm{B}) - 1 \right]. \tag{17}$$

Note that (13)-(17) describe the *temporal* growth and saturation of the instability. This form is suitable for comparison with observations of temporal growth, such as those of Clanet and Searby [17]. However we prefer here to make a comparison with the *spatio-temporal* observations of Truffaut and Searby [22], where the images are of better quality. We will again suppose, as in section 3, that the growth rate of the instability is sufficiently slow that the physical time t can be replaced by a Lagrangian time x/u_t, where u_t is the convection velocity of the structures. Equation (15) then takes the form [31] :

$$\alpha(x,t) = -A\,(x/u_t) - \frac{2\Omega}{a\, k_n} \log \left[1 - \frac{\cos(k(x - u_t t))}{\cosh\{k\mathrm{B}\,(x/u_t)\}} \right]. \tag{18}$$

Local maxima and minima of (18) occur when $\cos(k(x - u_t t)) = \pm 1$. The peak to peak amplitude of the structures, $\alpha_{p-p}(x)$ is thus easily found :

$$\alpha_{p-p}(x) = \frac{2\Omega}{a\, k_n} \log \left[\frac{\cosh\{k\mathrm{B}\,(x/u_t)\} + 1}{\cosh\{k\mathrm{B}\,(x/u_t)\} - 1} \right]. \tag{19}$$

This result can easily be compared with the experimental observations in Fig.10. The unknown quantities in (19) and (17) are $a(E)$, $\Omega(E)$ and k_n. The parameters $a(E)$ and $\Omega(E)$ were evaluated using the expansion coefficient in table 1. The parameter k_n was obtained from the experimental dispersion relation, Figs.12, by adjusting a parabola $\sigma\tau_t = \Omega k/k_n(1 - k/k_n)$ through the measured points.

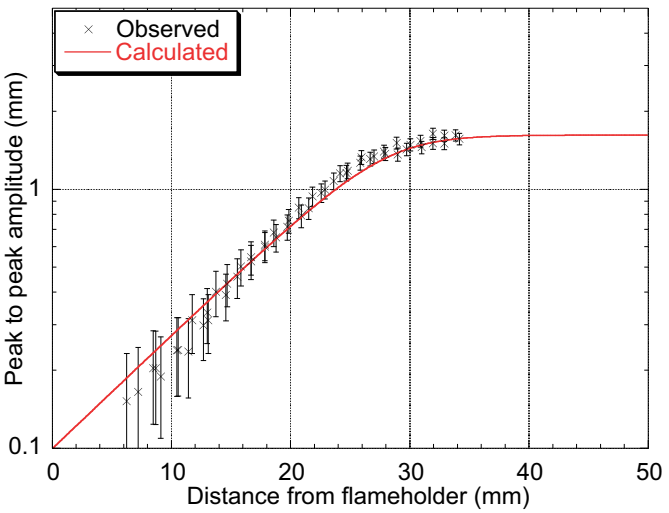

Fig. 14. Comparison of prediction of (19) with the data from Fig10. $\Omega(E) = 1.79$, $a(E) = 1.78$, $k/k_n = 0.38$

The mean tangential flow velocity, u_t, was 8.59 m/s. The result of the calculation is shown in Fig.14, along with the experimental points taken from Fig.10. The only adjustable parameter in this plot is the initial amplitude of wrinkling. All the other parameters are predetermined. It is not really surprising that the exponential part of the growth is well represented, this just means that, to order k^2, the Michelson–Sivashinsky equation has the same dispersion relation as the Clavin–Garcia equation. It is however remarkable that the Michelson–Sivashinsky equation also correctly predicts both the saturation amplitude of the structures and the width of the cross-over region between exponential growth and constant saturated amplitude.

The complete non-linear flame profile is given by (18). We have used this equation to calculate the curve shown in Fig.15. We have used the same input data as in Fig.14. Extra experimental parameters are the initial flame position and the initial phase of wrinkling. The excitation frequency is 2100 Hz. As expected from Fig.14, (18) gives a good overall prediction of the growth of the instability. It also gives a good representation of the form of the wrinkling, evolving from a sinusoidal to a cusped shape. The precision of the measurements of the flame profile from Fig.8 is insufficient to make any direct comparison concerning the radius of curvature of the cusps. However it is clear from the comparison of the observed and calculated flame profiles in Fig.15, that the flame front accelerates as the wrinkling increases and the amplitude of the acceleration is strongly underestimated by (18), which predicts that the increase in average flame speed should be proportional to the increase in flame area.

The reason for this strong increase in overall flame propagation speed is not well understood at this time. It is possible that the increase in average gas flow through the flame, arising from the increased amplitude of wrinkling with

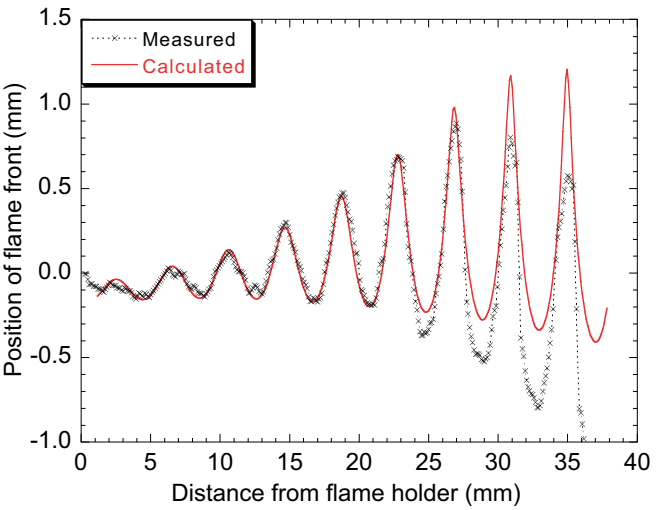

Fig. 15. Comparison of prediction of (18) with the flame profile of Fig.9. Input data as in Fig.14.

the Lagrangian coordinate, is giving rise to a large scale bending of the global hydrodynamic flow.

5 Conclusion

We have given a brief qualitative description of the mechanisms controlling the stability of planar laminar premixed flames, along with a historical survey of the theoretical approaches used to model them analytically. Despite the fact that the intrinsic instability of planar flames has been recognised for nearly a century and that the first attempts at an analytical description were performed over fifty years ago, experimental validation of the predicted growth rates of unstable structures has been obtained only recently. The main reason for this lies in the experimental difficulty of preparing and controlling an initially planar premixed flame front in a regime where the planar front is unstable.

We have described two recent experiments in which the growth rate of cellular structures on planar flames has been measured directly. In the first experiment an unstable laminar flame front is maintained initially planar by the action of an imposed acoustic field. The temporal growth rate of 2-D wrinkling is then observed after removal of the stabilising acoustic field. In the second experiment the spatio-temporal growth of wrinkling is observed on an inclined anchored flame. The residual turbulence of the gas flow is sufficiently low that turbulence induced perturbations at the base of the flame are not amplified to a perceptible amplitude before reaching the tip of the flame. Perturbations of controlled wavelength and amplitude are then created by an electrostatic deflection system. This technique has permitted the exploration of a significant portion of the

dispersion curve. The spatial growth rate of the wrinkling was converted to a temporal growth rate using a Lagrangian time. These experimental observations have confirmed the validity of the analytical theory.

In the final section of this paper we have investigated the non-linear saturation of the instability. Our experimental observations have been compared to the predictions of a Michelson-Sivashinsky equation phenomenologically extended to the case of strong gas expansion. It was seen that the extended equation correctly describes both the amplitude of structures at saturation and also the transition region from exponential growth to saturation. However experimental observation suggests that the overall increase in velocity of the wrinkled flame front is greater than predicted. The reason for this may lie in the generation of large scale hydrodynamic effects, absent from the Michelson-Sivashinsky description.

Acknowledgements

The authors are indebted to G. Joulin for his helpful discussions and suggestions concerning nonlinear analysis. They also wish to thank J. Minelli and F. Abetino for their invaluable technical assistance. Part of this work was carried out with the financial support of Air Liquide Welding.

References

1. Ya. B. Zeldovich: *The Theory of Combustion and Detonation of Gases.* Akademiia Nauk SSSR, Moscow, 1944.
2. Ya.B. Zeldovich and G. I. Barenblatt: Theory of flame propagation. *Combustion and Flame*, **3**, 61–74, 1959.
3. G. I. Barenblatt, Ya. B. Zeldovich, and A.G. Istratov: On the diffusional thermal stability of laminar flames. *Prikl. C. Mekh. Fiz.*, **4**, 21, 1962.
4. G.I. Sivashinsky: Diffusional-thermal theory of cellular flames. *Combustion Science and Technology*, **15**, 137–146, 1977a.
5. G. Joulin and P. Clavin: Linear stability analysis of nonadiabatic flames : diffusional-thermal model. *Combustion and Flame*, **35**, 139–153, 1979.
6. G. Darrieus: Propagation d'un front de flamme. Unpublished work presented at *La Technique Moderne*, and at *Le Congrés de Mécanique Appliqueée (1945)*, 1938.
7. L. Landau: On the theory of slow combustion. *Acta Physicochimica URSS*, **19**, 77–85, 1944.
8. Ya.B. Zeldovich, G.I. Barenblatt, V.B. Librovich, and G.M. Makhviladze: *The Mathematical Theory of Combustion and Explosions.* Plenum, New York, 1985.
9. G.H. Markstein: *Nonsteady flame propagation.* Pergamon Press, New York, 1964.
10. G.I. Sivashinsky: Nonlinear analysis of hydrodynamic instability in laminar flames. *Acta Astronautica*, **4**, 1177–1206, 1977b.
11. P. Clavin and F.A. Williams: Effects of molecular diffusion and of thermal expansion on the structure and dynamics of premixed flames in turbulent flows of large scale and low intensity. *Journal of Fluid Mechanics*, **116**, 251–282, 1982.
12. P. Pelcé and P. Clavin: Influence of hydrodynamics and diffusion upon the stability limits of laminar premixed flames. *Journal of Fluid Mechanics*, **124**, 219–237, 1982.

13. Paul Clavin: Dynamic behaviour of premixed flame fronts in laminar and turbulent flows. *Progress in Energy and Combustion Science*, **11**, 1–59, 1985.
14. P. Clavin and P. Garcia: The influence of the temperature dependence of diffusivities on the dynamics of flame fronts. *Journal de Mécanique Théorique et Appliquée*, **2**, 245–263, 1983.
15. R.J. Kee, F.M. Rupley, and J.A. Miller: The Chemkin thermodynamic data base. Report SAND87-8215B, Sandia National Laboratories, 1990.
16. G. Searby and J. Quinard: Direct and indirect measurements of Markstein numbers of premixed flames. *Combustion and Flame*, **82**, 298–311, 1990.
17. C. Clanet and G. Searby: First experimental study of the Darrieus-Landau instability. *Physical Review Letters*, **27**, 3867–3870, 1998.
18. Geoffrey Searby and Daniel Rochwerger: A parametric acoustic instability in premixed flames. *Journal of Fluid Mechanics*, **231**, 529–543, 1991.
19. N.W. McLachlan: *Theory and Application of Mathieu Functions*. Clarendon, Oxford, 1951.
20. I. Yamaoka and H. Tsuji: Determination of burning velocity using counterflow flames. *Proceedings of the Combustion Institute*, **20**, 1883–1892, 1984.
21. L.K. Tseng, M.A. Ismail, and G.M. Faeth: Laminar burning velocities and Markstein numbers of hydrocarbon/air flames. *Combustion and Flame*, **95**, 410–426, 1993.
22. J.M. Truffaut and G. Searby: Experimental study of the Darrieus-Landau instability on an inverted-'V' flame, and measurement of the Markstein number. *Combustion Science and Technology*, **149**, 35–52, 1999.
23. P. Garcia and R. Borghi: Etude de la stabilité de flammes prémélangées obliques. *Journal of Theoretical and Applied Mechanics*, Special Issue, 157–172, 1986.
24. M.L. Polanyi and G.H. Markstein: Phenomena in electrically and acoustically disturbed bunsen burner flames. Project SQUID technical report 5, Cornell Aeronautical Laboratory, 1947.
25. K.G. Payne and F.G. Weinberg: A preliminary investigation of field-induced ion movement in flame gases and its applications. *Proceedings of the Royal Society London*, **A250**, 316–336, 1958.
26. D. Bradley: The effects of electric fields on combustion processes. In *Advanced Combustion Methods*, pages 331–394. Academic Press, London, 1986.
27. S. Kwon, L.-K. Tseng, and G.M. Faeth: Laminar burning velocities and transition to unstable flames in $H_2/O_2/N_2$ and $C_3H_8/O_2/N_2$ mixtures. *Combustion and Flame*, **90**, 230–246, 1992.
28. D.M. Michelson and G.I. Sivashinsky: Nonlinear analysis of hydrodynamic instability in laminar flames. *Acta Astronautica*, **4**, 1207–1221, 1977.
29. Guy Joulin and Pierre Cambray: On a tentative approximate evolution equation for markedly wrinkled premixed flames. *Combustion Science and Technology*, **81**, 243–256, 1992.
30. O. Thual, U. Frisch, and M. Henon: Application of pole decomposition to an equation governing the dynamics of wrinkled flame fronts. *Journal de Physique*, **46**, 1485–1494, 1985.
31. Guy Joulin: Private communication, 1999.

Gaseous Detonations

Paul Clavin

[1] Institut de Recherche sur les Phénomènes Hors Équilibre
UMR 6594. CNRS-Universités d'Aix-Marseille I&II
S. 252 Centre Universitaire de St Jérôme, 13397 Marseille Cedex 20, France
[2] Institut Universitaire de France.

Abstract. Recent theoretical results on gaseous detonations are presented, including initiation, one-dimensional nonlinear oscillations (galloping detonations) and multidimensional cellular structures.

1 Introduction

Combustion of premixed gases is a typical example of an irreversible phenomenon occurring far from chemical equilibrium. The occurrence of a globally exothermic reaction requires a very large number of molecular collisions. The rate of heat release is thus much slower than the relaxation toward the local equilibrium described by the Maxwell-Boltzmann distribution. A macroscopic description is then valid and the complete set of governing equations may easily be written down. Nevertheless, the detailed chemistry is quite complex and many elementary steps are still not well known. The full complexity is not needed to describe many of the aspects of combustion.

Nineteenth century experiments on combustion waves led to the distinction between deflagrations and detonations. The main features of these waves may be represented by a simple model relating the overall heat release rate, $1/\tau_r$, to the elastic collision frequency, $1/\tau_c$, through an Arrhenius factor with an activation energy larger than the thermal energy,

$$\frac{1}{\tau_r} = \frac{1}{\tau_c} e^{-\frac{E}{RT}} \quad with \quad \frac{E}{RT} >> 1, \tag{1}$$

where, for simplicity, a prefactor of order unity is incorporated in $1/\tau_c$. Notice that the temperature is defined only when the frequency of elastic collisions (mechanism restoring the local Maxwell-Boltzmann distribution) is higher the rate of heat release (inelastic collisions). Therefore equation (1) is meaningful only for an activation energy which is sufficiently large as compared to the thermal energy. The high sensitivity to temperature in the Arrhenius factor then explains the existence of a cold mixture chemically frozen far from equilibrium, a necessary ingredient for premixed combustion.

In planar flames the steady propagation results from a balance between heat diffusion and reaction rate. A dimensional analysis then yields a propagation velocity in the form $U_L \approx \sqrt{D_T/\tau_r}$ where D_T is the heat diffusion coefficient. According to kinetic theory, the diffusion coefficients in gases are almost all

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 182–189, 2001.
© Springer-Verlag Berlin Heidelberg 2001

equal and of order $D \approx a^2\tau_c$ where a is the speed of sound. Equation (1) then shows that the propagation Mach number is small, $U_L/a \approx e^{-E/2RT} << 1$, and the flame is quasi-isobaric. Mass conservation then implies that the relative variations of flow velocity, specific volume and temperature (across a steady and planar flame) are equal, $U_b/U_L = \rho_u/\rho_b = T_b/T_u$, where subscripts u and b denote the unburned and burned mixture, respectively. The laminar flame speed appears as an eigenvalue of a nonlinear reaction-diffusion problem solved in the singular limit $\beta \equiv E(T_b - T_u)/RT_b^2 \to \infty$ which captures the specificity of flames when compared to ordinary reaction-diffusion waves [1-3].

Detonations are combustion waves propagating at a supersonic speed (hyperbolic problems). Due to the large activation energy of the reaction rate, see equation (1), and to a large heat release (compared to the thermal enthalpy), the inner structure of a 1-D detonation wave propagating at constant speed is, according to Zeldovich-Neumann-Döring (ZND), constituted by a strong inert shock followed by a subsonic reacting flow [4]. The flow velocity relative to the shock is sufficiently large for the diffusion mechanisms to be negligible.

2 Initiation

There are different ways to initiate a detonation, including transitions from deflagration to detonation [5]. This mechanism, however important it is for practical purpose, will not be addressed here. The details of the onset of detonation by auto-explosion centers have been revealed thirty years ago [6]. However self ignition of a uniform pocket of reactive mixtures produces shock waves which are too weak to initiate a detonation. An early numerical analysis showed that a detonation might be spontaneously formed in a gas at high temperature by a non uniform temperature distribution [7]. This phenomenon was explained some years later [8][9] and may be understood from the wave equation of sound propagating in a burning medium,

$$\partial^2 p/\partial^2 t - a^2 \Delta p = \partial \ddot{q}/\partial t, \ddot{q} \equiv \ddot{Q}(\underline{r}, t)(C_p - C_v)/C_v \tag{2}$$

in which inhomogeneities of the flow have been omitted for simplicity and $\ddot{Q}(\underline{r}, t)$ is the heat release rate per unit volume. An initial temperature gradient in the fresh mixture produces a gradient of induction time $\tau_I(x)$ which, in turn, produces an ignition front which propagates at a speed $\frac{1}{d\tau_I/dx}$. When this speed is equal to the sound speed, the heat release is synchronized with the pressure waves, leading very quickly to a pressure overshoot, sufficient to initiate a Chapman-Jouguet detonation (CJ wave). However more recent studies [10][11] have shown that additional conditions are required for the transmission of a detonation into a cold mixture: The same gradient responsible for a spontaneous formation of the detonation at high temperatures may well also quench the detonation at lower temperatures.

Before getting a physical insight into the quenching mechanism it is worth discussing the direct initiation of gaseous detonations by an energy source, a

problem which was first studied during the 1950s [12]. A detailed bibliography may be found in reference [13]. A strong blast wave is first initiated by the energy deposition, viewed as an ideal point source at which the energy is instantaneously deposited. The strength of the initial blast wave is an increasing function of the energy. Below a critical energy, the strongly overdriven detonation which is first established decays rapidly and the reaction front separates from the leading shock. At last a premixed flame trails behind the inert shock and no detonation is initiated in the cold mixture [14]. For energies larger than the critical threshold, the strongly overdriven detonation relaxes to an expanding spherical CJ wave. A rough criterion was first proposed by evaluating the critical energy from both the energy per unit volume Q liberated by the chemical reaction and the thickness of the CJ detonation wave l_{CJ} which is controlled by the chemical kinetics [12]. The so obtained result, Ql_{CJ}^3 , underestimates the critical energy by a factor between 10^6 and 10^8 ! This indicates that the characteristic length must be at least one hundred times larger than l_{CJ}. An explanation of the phenomenon presented below is provided in reference [13].

Detonation initiations are fully unsteady problems that may be investigated, in a first approximation, by theoretical analyses using 1-D geometry. The inner structure of the wave is perturbed both by unsteadiness and curvature of the front. The critical conditions for sharp transitions, such as detonation quenching by gradients or initiation failure, are mainly controlled by quasi-steady geometrical effects [10][11][13]. The main features may then be described by an approximated analysis neglecting unsteadiness involving short time scales, of the order of the transit time across the detonation structure. A one-step Arrhenius law in the limit of large activation energy is then the simplest convenient model. In this model the inner structure of a detonation wave reduces to the square-wave model in which all the heat is instantaneously released after a temperature-dependent induction time, $l_{CJ}(T_N)$, where T_N is the temperature of compressed gas at the Neumann spike, just behind the inert shock. Following Zeldovich, nonlinear quasi-steady 1-D problems are easily solved analytically with such a piecewise constant (square-wave) model [2]. The induction length $l_{CJ}(T_N)$ is very sensitive to T_N (and thus to variations of the propagation velocity). This strong non linearity is responsible for the sharp transitions observed in experiments: As a rule, the CJ propagation speed is a solution of the nonlinear equation which presents a turning point (in parameter space) corresponding to a critical conditions, a phenomenon first observed by Zeldovich in his study of critical tube radius [2][4].

In spherical geometry, the front curvature changes the inner structure of quasi-steady detonations. For direct initiation of detonations, a non linear curvature effect is responsible for initiation failure [13]: Below a critical radius there is no longer a marginal solution of the CJ type (lower bound of propagation velocity) and the flow of the burned gas is subsonic relative to the shock, however small the propagation speed may be. The leading shock is then slowed down without limit by the rarefaction wave in the burned gas flow. A CJ detonation can be initiated only if the initial energy is sufficiently large, so that the prop-

agation speed of the overdriven detonation (triggered by the initial blast wave) approaches from above the CJ detonation speed with a radius of the front larger than the critical radius. The critical radius is found to be larger than the planar CJ thickness by a factor $16e(E/RT_N)C_P^2/(C_P^2 - C_V^2)$ which is of order 10^2. The corresponding criterion is in good agreement with 1-D numerical simulations showing that the unsteady effects (not taken into account in a quasi-steady analysis) just increase the critical energy by a factor of order unity, not larger than 3.

The quenching by a temperature gradient in the fresh mixture is more subtle [10][11]. The sensitivity of the induction length to temperature makes the problem unsteady: The variations of propagation velocity of the inert shock and of the reaction zone are different. However, for small gradients, the characteristic time scale is larger than the transit time and the "slow evolution" may be described by a quasi-steady approximation The two Michelson-Rayleigh straight lines which intersect at the Neumann state in the pressure-volume diagram have different slopes. The Michelson-Rayleigh line of the inert shock starts from the initial condition in the fresh mixture with a smaller slope than the steady CJ wave. In order to satisfy a sonic condition in the burned gas (imposed by the rarefaction wave), the Michelson-Rayleigh line of the reaction zone (starting from the same Neumann state) is tangent to the Hugoniot curve in the burned gas. For a given temperature gradient in the fresh mixture, the "shock-reaction complex" is a solution of the nonlinear equation which presents a turning point at a critical value of the gradient [10]. For a non dimensional gradient (reduced with the local CJ wave thickness) larger than the critical value there is no quasi-steady solution. Numerical results of 1-D unsteady detonations show that quenching occurs precisely at the critical conditions predicted by the quasi-steady analysis [10][11].

The quasi-steady analyses presented above implicitly assume that the intrinsic dynamics of detonation waves are stable with a relaxation time shorter than the characteristic evolution time of the quasi-steady solutions. We will now see in the next section that this is not generally the case and that the intrinsic dynamics of planar detonation waves cannot be studied with the same model.

3 Galloping Detonations

One-dimensional nonlinear oscillations of planar detonations, so-called galloping detonations, were predicted by numerical simulations in the 1960s and observed in experiments a few year later, see references [15][16] for a detailed bibliography. As shown in these references, the phenomenon is not related to a thermo-acoustic instability, as it was thought. This is clearly exhibited by the study of the limiting case of overdriven detonations in the Newtonian approximation, $(\frac{C_P}{C_V} - 1) << 1$, yielding a quasi-isobaric heat release. Heating by adiabatic compression is then negligible and the distribution of heat release rate per unit length, $w(x,t)$, is mainly controlled by the entropy wave which propagates downstream from the inert shock to the burned gas. In the referential frame of the perturbed shock,

equations for temperature and species then form an autonomous system,

$$C_p DT/D\tau = Q\Omega(..Y_j..,T), \quad DY_i/D\tau = \Omega_i(..Y_j..,T), \tag{3}$$

where Y_i is the mass fraction of species j, the rates $\Omega(..Y_j..,T)$ and $\Omega_i(..Y_j..,T)$ being prescribed by the chemistry. When fluctuations of the shock velocity are relatively small, the material derivative may be written in terms of non dimensional time and mass weighted distance from the leading shock (τ , ξ) as $D/D\tau = \partial/\partial\tau + \partial/\partial\xi$. Boundary conditions at the shock, $\xi = 0$, are given by the composition in the fresh mixture and the temperature at the Neumann state of the perturbed shock $T_N(\tau)$. Solution of equation (3) then yields the instantaneous distribution of heat release rate, $w(\xi,t)$, in terms of the history of the Neumann temperature as $w(\xi,t) = W(\xi,\, T_N(\tau-\xi))$, where the time lag results from the propagation of the entropy wave, $W(\xi,\, T_N)$ being the steady state solution corresponding to a fixed temperature T_N at the Neumann state. Assuming that these distributions are very sensitive to T_N, it is convenient to introduce a large non dimensional parameter, $\beta >> 1$, describing the sensitivity to T_N by rewriting $W(\xi,T_N)$ in the form $\mathrm{W}(\xi,\theta_N)$ with $\theta_N \equiv \beta(T_N - T_{No})/T_{No}$ where T_{No} is the Neumann temperature of a reference state and where $\partial W(\xi,T_N)/\partial T_N = \beta\partial \mathrm{W}(\xi,\theta_N)/\partial\theta_N$, with $\partial \mathrm{W}(\xi,\theta_N)/\partial\theta_N = O(1)$. Solution of equation (3) then yields $w(\xi,t) = \mathrm{W}(\xi,\, \theta_N(\tau-\xi))$ showing how a small temperature variation at the Neumann state, $\delta T_N(\tau)/T_{No} = O(1/\beta)$, (i.e. small relative variations of the shock velocity) produces strong quasi-isobaric modifications to the inner detonation structure. For fast detonations in the Newtonian approximation, the Mach number of the flow (relative to the shock) is sufficiently small to ensure the validity of a quasi-isobaric approximation throughout the inner structure of detonations at strong overdrive. Mass conservation then yields a nonlinear integral equation which is obtained by integrating the continuity equation from the perturbed shock (where the Rankine-Hugoniot jump conditions hold) to the piston located far downstream in the burned gas [30],

$$b\theta_N(\tau) = \int_0^\infty [W(\xi,\theta_N(\tau-\xi)) - W(\xi,\theta_N=0)]d\xi, \tag{4}$$

with $\frac{1}{b} \equiv \beta\frac{C_P - C_V}{C_V}\frac{Q}{C_P T_{No}}$ yielding, in the linear approximation

$$b\delta\theta_N(\tau) = \int_0^\infty w_\theta(\xi)\delta\theta_N(\tau-\xi)d\xi, \;\; with \;\; w_\theta(\xi) \equiv [\partial W(\xi,\theta_N)/\partial\theta_N]_{\theta_N=0}\,. \tag{5}$$

The function $w_\theta(\xi)$ represents the deformation of the steady state distribution of heat release rate for slightly different values of the Neumann temperature. The parameter $1/b$ represents the strength of sensitivity to temperature. In equations (4) and (5), the only assumption concerning the chemistry is the high sensitivity of the steady distribution $W(\xi,\, T_N)$ to T_N. Solutions of integral

equations (4) and (5) are asymptotic solutions of the reactive Euler equations describing galloping detonations (1-D unsteady).

For a given distribution $w_\theta(\xi)$, assumed to not be singular, a Landau-Hopf bifurcation appears in the solution of equations (4) and (5) when $1/b$ is increased, describing the onset of an oscillatory instability. Smaller critical values of $1/b$ are obtained with stiffer distributions $w_\theta(\xi)$. The instability is enhanced by increasing the stiffness of $w_\theta(\xi)$ as well as the values of β and Q. In the limit of a large activation energy (with a reduced heat release of order unity) the one-step Arrhenius law yields the square-wave model: $w_\theta(\xi) = \delta(\xi - 1)$. In this limit, equation (5) degenerates to a non physical difference-differential equation of the advanced type, $b\,[\theta_N(\tau) - 1] = d\theta_N(\tau - 1)/d\tau$, yielding a singular dynamics of the wave. Such a model, so useful in flame theory, is thus not convenient for describing the onset of the oscillatory instability of planar detonations.

Small effects of acoustic waves have also been incorporated at moderate overdrive [16]. The presence of acoustic waves introduces a stabilizing factor. This confirms that the triggering mechanism of galloping detonations is not a thermoacoustic instability reviewed in reference [17] and that strong overdrive is a convenient limit to exhibit the essential phenomena.

4 Cellular Fronts

Explanations of the transverse instability of overdriven detonations have also been provided in the limiting case of the preceding section [18]. Physical insights are more easily gained in this case because quasi-isobaric expansion and compressibility of the reacting gas are now coupled in a simple way. An oscillatory instability of the inner structure of wrinkled fronts of overdriven detonations is produced by flow-induced perturbations of the quasi-isobaric combustion zone. The oscillatory nature is related to underlying oscillatory modes of the inert shock, first described in the Russian literature of the1950s, see reference [18] for a detailed bibliography. The instability develops even without sensitivity to temperature and presents an oscillation period of the order of the transit time of a fluid particle across the inner structure. Acoustic waves in the burned gas damp out the perturbations with small wavelengths which would be otherwise amplified by the quasi-isobaric instability of the reaction zone. For similar reasons to those presented in the preceding section, asymptotic analysis leading to the square-wave model is not convenient to study the multidimensional dynamics of detonation waves. The early analysis of Erpenbeck [20-22] found the stability limits and identified the conditions at the instability threshold: a small heat release and a smooth distribution of the rate of heat release. Following the methodology used by Sinvashinsky [23][24] for flames in the approximation of a small gas expansion, an amplitude equation for the nonlinear evolution of weakly unstable detonations was obtained in a systematic way by a bifurcation method [19], valid for weakly unstable detonations at strong overdrive,

$$\frac{\partial^2}{\partial\tau^2}\alpha(\tau,\underline{\eta}) - a^2\Delta\alpha(\tau,\underline{\eta}) + \frac{\partial}{\partial\tau}\left|\underline{\nabla}\alpha(\tau,\underline{\eta})\right|^2 = -2(\epsilon\sqrt{q})\frac{\partial}{\partial\tau}L(\alpha)$$

$$+q\frac{\partial^2}{\partial\tau^2}\int_0^\infty E(\xi)\alpha(\tau-\xi,\underline{\eta})d\xi+\Delta\int_0^\infty F(\xi)\alpha(\tau-\xi,\underline{\eta})d\xi, \tag{6}$$

where $\xi=\alpha(\tau,\underline{\eta})$ is the equation of the shock surface ξ and $\underline{\eta}$ being the longitudinal and the transverse non dimensional coordinates. The gradient operator $\underline{\nabla}$ refers to the non dimensional transverse coordinate $\underline{\eta}$, a is the sound speed in the compressed gas, q is the non dimensional heat release $q=Q/C_PT_{No}$, ϵ is the Mach number at the Neumann state which is a small parameter in the limiting case under investigation. Finally $E(\xi)$ and $F(\xi)$ are functions which characterize the steady state distribution of heat release,

$$E(\xi)\equiv\beta\frac{C_P-C_V}{C_V}w_\theta(\xi),F(\xi)\equiv w_o(\xi)+\frac{\partial}{\partial\xi}\left[\xi w_o(\xi)\right] \tag{7}$$

with $w_o(\xi)\equiv W(\xi,\theta_N=0)$, and $L(\alpha)$ is a linear term coming from the acoustic waves in the burned gas and the linear operator $L(.)$ acts on functions depending on the transverse coordinates $\underline{\eta}$, and is defined in the Fourier space by a multiplication by the modulus of the wave vector. According to the minus sign in front of it, this term stabilizes transverse perturbations with small wavelengths. The two last terms in the r.h.s. of equation (6) represent the quasi-isobaric linear instability which develops in 2 or 3-D geometry even when $\beta=0$. The two first terms in the l.h.s. represent a wave equation describing the neutral oscillatory modes of the inert shock. The third term in the l.h.s is the dominant nonlinear term. It is not of a geometric nature (as it was for weakly unstable flames) but comes from the non linear flow of the entropy-vorticity wave. This term is responsible for the Mach stem formation. Equation (6) is valid in the Newtonian approximation for overdriven detonations close to the instability threshold, $\epsilon<<1$, $(C_P-C_V)/C_V=O(\epsilon^2)$, $q=O(\epsilon^2)$, $\beta(C_P-C_V)/C_V=O(1)$ or smaller. In the planar geometry, equation (6) reduces to (5). Preliminary results of the numerical analysis of equation (6) show a qualitative agreement with patterns observed in experiments. However, due to the approximations used in (6), a quantitative agreement with the fully established multi-headed CJ detonations cannot be expected.

5 Conclusions

Theoretical results obtained by systematic perturbation methods may be relevant, at least qualitatively, even when the underlying approximations are not quite well verified. The main task is to pick up the key mechanisms in a sufficiently simple form to be described analytically at the leading order of a perturbation analysis. Generally speaking, this is the case for the "amplitude equations" concerning weakly unstable systems and obtained by a bifurcation method whose mathematical validity is limited to the vicinity of instability onset. Surprisingly, these results are often proved to be still relevant even for strongly unstable regimes, as it was illustrated for flames during the two last decades.

This has been recently achieved for the cellular fronts of overdriven detonations by using a nonlinear perturbation analysis around the linear stability limits.

The cellular structure results from an oscillatory instability of the internal structure of the detonation front. This instability is associated with the entropy-vorticity waves and may be described within the framework of a quasi-isobaric approximation valid for overdriven detonations when the difference of specific heat is small. Much still remains to be done along this line to incorporate the cellular fronts of CJ detonations waves.

References

1. Zeldovich, Ya. B., and Frank-Kamenetskii, D., *Azhur. Fiz. Khim.* 12, 1:100-105 (1938).
2. Ostriker, J.P., *Selected works of Y.B. Zeldovich*, Vol. I. Princeton Univ. Press, 1991.
3. Clavin, P. and Liñan, A., *NATO ASI Series* B: Physics Vol 116, Plenum Press, 1984, pp. 291-338.
4. Zeldovich, Ya.B. and Kompaneets, A.S., *Theory of detonation*, Academic Press, 1960.
5. Lee, J.H.S. and Moen, I.O., *Prog. Energy Combust. Sci,* 6:359-389 (1980).
6. Urtiew, P.A. and Oppenheim, A.K., *Eleventh Symposium (International) on Combustion*, The Combustion Institute, Pittsburg, 1969, pp. 853-864.
7. Zeldovich, Ya.B., Librovich, V.B., Makhvildaze, G.M. and Sivashinsky, G.I., *Astronautica Acta*. 15: 313-321(1970).
8. Lee, J.H., Knystautas, R., and Yoshikawa, N., *Acta Astronautica* 5:971-982 (1978).
9. Zeldovich, Ya.B., *Combust. Flame* 39:211-214 (1980).
10. He, L. and Clavin, P., *Twenty-Fourth Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh PA, 1992, pp.1861-1867.
11. He, L. and Clavin, P., *Twenty-Fifth Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh PA, 1994, pp.45-51.
12. Zeldovich, Ya.B., Kogarko, S.M. and Simonov, N.N., *Sov. Phys. Tech. Phys.* 1(8):1689- 1713 (1956).
13. He, L. and Clavin, P., *J. Fluid Mech.* 277:227-248 (1994).
14. Lee, J.H., Knystautas, R. and Bach, G.G., "Theory of explosions" 1969 McGill University report ASFOR 69-3090 TR.
15. Clavin, P. and He, L., *J. Fluid Mech.* 306:353-378 (1996).
16. Clavin, P. and He, L., *Phys. Rev. E* 53, 5:4778-84 (1996).
17. Clavin, P., *Ann. Rev. Fluid Mech.* 26:321-352 (1994).
18. Clavin, P., He, L. and Williams, F. A., Phys. Fluids 9,12:3764-3785 (1997).
19. He, L. and Clavin, P., *J. Fluid Mech.* Submitted (2000).
20. Erpenbeck, J.J., *Phys. Fluid* 7(5):684-696 (1964).
21. Erpenbeck, J.J., *Phys. Fluid* 9:1293-1306 (1966).
22. Erpenbeck, J.J., *Twelfth Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh PA, 1969, pp. 711-721.
23. Sivashinsky, G. I., *Acta Astronaut.* 4:1177-1206 (1977).
24. Sivashinsky, G. I., *Ann. Rev. Fluid Mech.* 15:179-199 (1983).

The Vortical Structure of Flame Spreading over Liquid Fuels

Francisco J. Higuera[1], José C. Antoranz[2], Vladimir Sankovitch[2], José L. Castillo[2], Eugenio Degroote[3], and Pedro L. Garcia-Ybarra[4,2]

[1] ETSI Aeronáuticos, Univ. Politécnica de Madrid, Spain
[2] Dept. Física Matemática y Fluidos, UNED, Madrid, Spain
[3] EUIT Agrícola, Univ. Politécnica de Madrid, Spain
[4] Dept. Combustibles Fósiles, CIEMAT, Madrid, Spain

1 Introduction

Understanding flame spreading is a matter of both fundamental interest and crucial practical importance, mainly for its relevance to safety issues, as it is the base for the knowledge and ultimate control of fire propagation/suppression mechanisms. The fuel vapor surrounding a condensed fuel in an oxidizing atmosphere can support the propagation of a flame over the fuel surface. If the fuel temperature, and therefore the fuel vapor pressure, is high enough, then propagation occurs in the premixed (triple) flame regime [1] through the fuel vapor-air mixture. If the fuel temperature is decreased down to values that make this combustion regime impossible, due to heat loses towards the fuel and/or to adverse kinetics conditions, then the chemical reaction can still proceed, provided the chemical heat release is high enough to vaporize such an amount of fuel as to increase locally the Damkholer number above its extinction value. In this case, flame spreading relies on relatively slow mechanisms of heat and mass transfer across the fuel-gas interface, which render spreading slower than it was at high temperatures. When the fuel is a liquid, these mechanisms include convection, induced by thermocapillarity and/or buoyancy, which leads to characteristic propagation regimes [2] absent for solid fuels; see [3] for a review. Thermocapillary convection around a fluid-fluid interface occurs when temperature differences along the interface establish an imbalance of surface forces by inducing changes in the surface tension. The fluid is then set into motion, generally down the surface temperature gradient. The importance of thermocapillary convection in flame spreading over liquid fuels was stressed by Sirignano and Glassman [4] and, despite the abundant literature devoted to this subject (see, for example, Akita [2]), several important questions remain. This paper reviews a series of experiments and studies [6,7,8,9,10] carried out to characterize flame spreading over liquid fuels of relatively simple molecular structure, so as to avoid complex vaporizing/combustion phenomena. The first four aliphatic alcohols (methanol, ethanol, propanol and butanol) have been selected to perform the experiments. In addition, buoyancy effects are weak compared to thermocapillary convection in the small pools used in this work, and can be left out of the analysis. Thus, the following discussion of experimental results and numerical simulations focuses

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 190–201, 2001.
© Springer-Verlag Berlin Heidelberg 2001

on the effects of thermocapillary convection on the global heat balance in the liquid, under the restrictions imposed by the boundary condition at the fuel surface required by the spreading flame. The experimental setup and procedures are described in Section 2. Section 3 contains a summary of the main results about flame spreading velocity measurements, while results of numerical simulations are presented in Section 4. Finally the last section is devoted to a discussion of the experimental results in the light of the numerics, and the main conclusions are exposed.

2 Experimental Setup and Results

Experiments were conducted in two open aluminum trays of different lengths (40 cm and 100 cm) with the same width (2.5 cm) and depth (4 cm). The temperature of the bottom of the trays was controlled by a thermal bath. The lateral walls were fitted with windows of Pyrex, 12 cm long and 1 cm height, to allow optical access to the liquid bulk. In each series of experiments, the tray was filled with an alcohol, flush with the upper border, and a flame was ignited at one end and left to spread freely along the channel. Then the flame was extinguished and the experiment repeated with a different bath temperature.

As sketched in Fig. 1, three different techniques were used simultaneously to characterize the flame spreading. First, a video camera allowed to measure the flame spreading velocity by recording a general sight of the progressing flame. Second, another camera recorded the vision through the lateral windows, lit up by a Teopler schlieren system which provides a visualization of the thermal gradients induced in the alcohol by the spreading flame. And third, an array of eight 25 μm thick thermocouples was immersed by the bottom of the tray and bent toward the surface in such a way that the eight welding points were disposed along the surface at intervals of 2 cm, allowing the measurement of fuel surface temperatures with minimal disturbance of the flame. Also, measurements of the passage time of the flame front between pairs of successive thermocouples were used as an alternative method of determining the spreading velocity, which proved more accurate than the video records for fast spreading regimes (about 100 cm/s). For moderate to low spreading regimes (from about 10 to 1 cm/s), the video records were accurate enough to determine with precision the observed variations of the spreading velocity. To this end, a procedure was deviced that generates a picture by pasting, one below the other, slits of consecutive video frames corresponding to the middle of the channel. The result was like a classical streak photograph, with the border between the light and dark regions marking the position of the flame tip, whose derivative is the flame spreading velocity. Three examples of these compositions are shown in Fig. 2. Case a) corresponds to uniform spreading, and cases b) and c) correspond to pulsating spreading with different pulsation periods.

The different spreading regimes characterized by these three techniques will be described and analyzed in the next section.

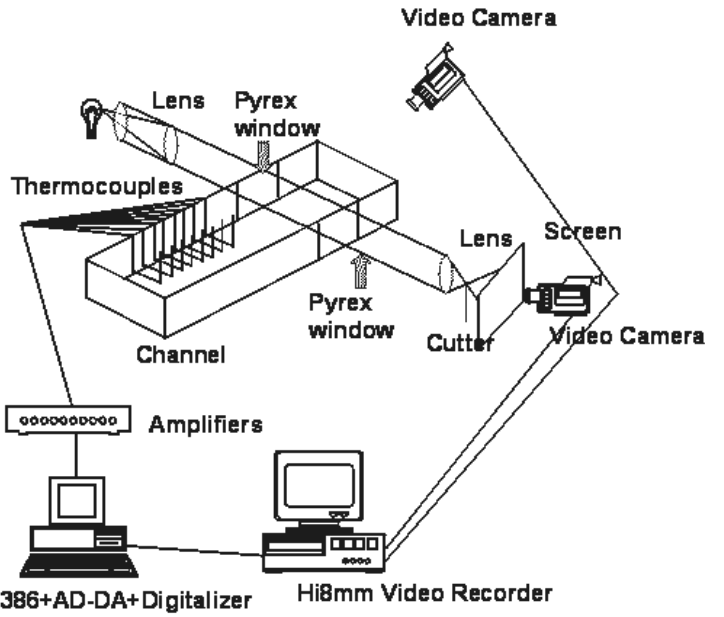

Fig. 1. Sketch of the experimental set up showing the three techniques used to characterize flame spreading: (1) recording of the upper view; (2) recording of a transverse schlieren view; and (3) records of eight thermocouples located at the alcohol surface.

3 Regimes of Flame Spreading at Different Temperatures

The thermal bath allows easy experimental control of the fuel temperature in the channel, which is therefore the most convenient control parameter to study the different spreading regimes. The measured flame spreading velocity at different fuel temperatures shows the same qualitative features for all the alcohols that were tested. The results for methanol, ethanol and propanol are shown in Fig. 3. The common features are summarized in Fig. 4, on the methanol curve, and described below.

The condensed fuel heated by the flame plays a dual role in the flame spreading process, being a sink of heat and a source of fuel vapor at the same time. At high temperatures, the fuel vapor pressure is sufficiently large for a flammable mixing layer of fuel vapor and air to develops in the gas phase above the fuel. Once a flame is ignited in this mixture, it propagates along the layer in the premixed regime, at speeds above 100 cm/s. The flame progresses with the typical triple-flame structure [1] and the fuel surface acts primarily as a sink of heat that quenches the rich-side trailing edge of the triple flame in the vicinity of the surface. The smaller the initial fuel temperature, the larger the heat losses and the width of the quench layer. Ultimately, the quench layer and the mixing layer become of comparable thickness and the triple flame ceases to exist.

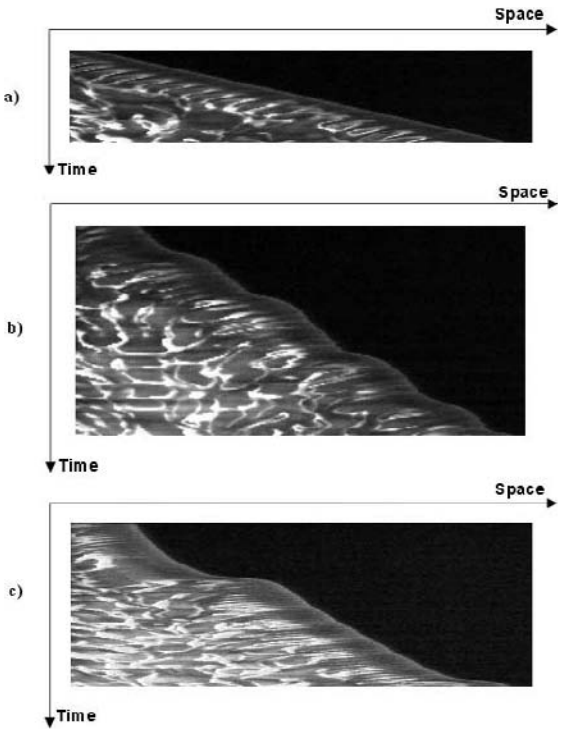

Fig. 2. Streak photographs of spreading flames. Case a) displays uniform spreading at relatively high fuel temperatures, with the flame progressing at constant velocity. In case b) temperature is below but close to the threshold temperature where pulsating spreading appears. In case c) the system is in the far-from-threshold pulsating regime where the spreading velocity features long period relaxation oscillations.

At this point, flame propagation switches to a creeping mode where both the heat losses towards the condensed fuel and the fuel vaporized become crucial to flame spreading. Under these conditions the flame spreading proceeds very much like in the case of a solid fuel and de Ris theory [11] can be applied to predict flame the spreading velocity [8,10]. According to this theory, the spreading velocity is determined by the balance between the heat flux leaving the flame region towards the liquid fuel surface and the heat flux that can be evacuated in the liquid side of the same region. The heat losses increase and the spreading velocity decreases by decreasing the liquid fuel temperature, as can be seen in the region of Fig. 4 labelled solid-like regime.

A steep change in the surface tension of the fuel liquid surface underneath the flame tip is brought about by the temperature increase from the initial fuel temperature, far ahead of the flame, to the fuel boiling temperature, reached below the flame. As a consequence, in a reference frame moving with the flame, the incoming liquid fuel decelerates in the vicinity of the surface on approaching the flame, due to the action of thermocapillary stresses. When the characteris-

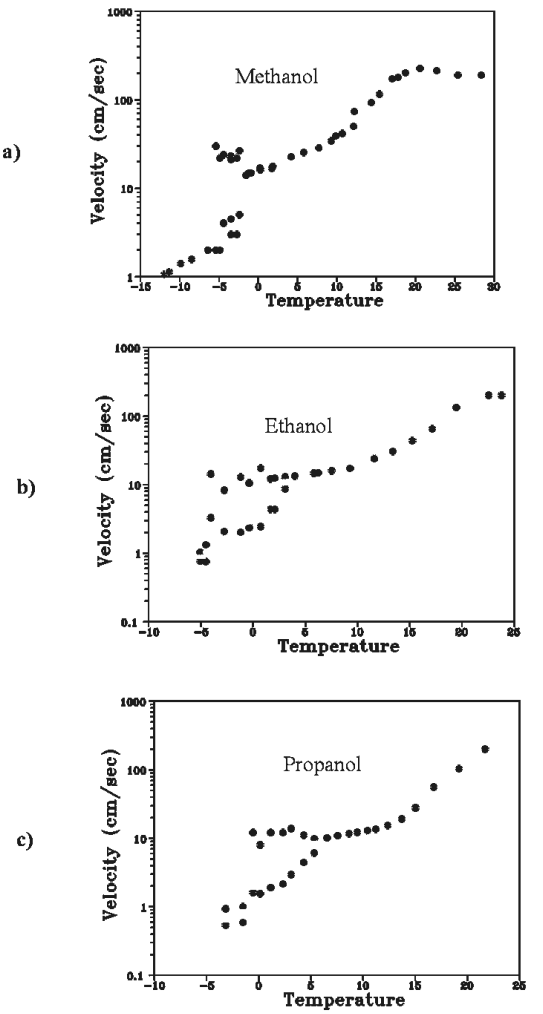

Fig. 3. Flame spreading velocity as a function of the fuel temperature for a) methanol; b) ethanol; and c) propanol.

tic velocity induced by the surface tension gradient becomes comparable to the spreading velocity, thermocapillary currents opposing the main stream appear in the region below the flame. Then a vortical structure develops which, by carrying warm liquid fuel ahead of the flame, reduces the heat losses and increases the spreading rate. This mode of flame spreading corresponds to the thermo-

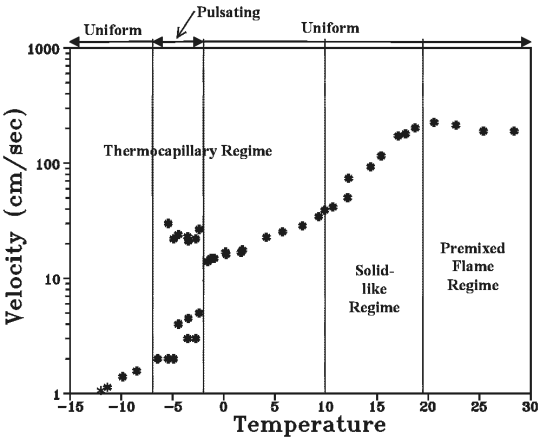

Fig. 4. Main characteristics of the different spreading regimes.

capillary regime of Fig. 4. The transition to this regime can be described by the value of the ratio between the characteristic thermocapillary velocity and the flame spreading velocity. Experimental data [10] point towards the existence of a critical threshold for this ratio, the same for all the alcohols, such that above it or, equivalently, for values of the fuel temperature below the corresponding threshold value, the spreading flame is preceded by a vortical structure. The temperature in this vortex, where warm and cold liquid fuel mix and recirculate, is typically 10^oC above the initial fuel temperature, and its size ranges from vanishing values at the transition point to several centimeters.

Pulsating spreading is obtained by further decreasing the liquid fuel temperature [2]. Fast and slow spreading rates alternate during the flame propagation along the channel. The spreading velocity, obtained from streak photographs, shows oscillations that evolve from the sinusoidal type, close to the threshold of this pulsating mode, to the relaxation type far from it, as can be seen in Fig. 5. Slightly away from the threshold, the growth of the amplitude and the period follows closely the well-known laws characterizing Hopf bifurcations [6,9]. The period increases as the relaxation oscillations become more pronounced, with long intervals of slow spreading and short spikes of high velocity, as in case c) of Fig. 5. Finally, the oscillations disappear abruptly, giving way to a new steady regime of uniform spreading.

The characteristics of this latter transition permit to classify it as a saddle-node infinite period (SNIPER) bifurcation [6], where the period diverges logarithmically as the limit cycle approaches a homoclinic trajectory. The node becomes the new spreading steady state, for which the length of the vortical structure preceding the flame reaches values of 10 to 20 cm and advances steadily with the flame, at rates of about 1 cm/s. A transverse view of this eddy, obtained

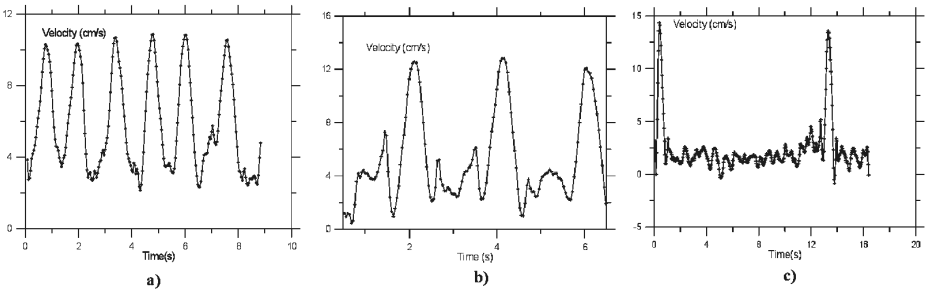

Fig. 5. Oscillations of the flame spreading velocity in the pulsating regime. a) Near-threshold. b) Far from threshold. c) Relaxation oscillations.

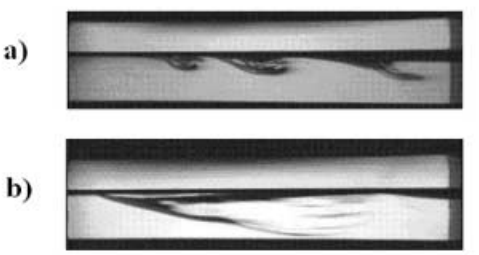

Fig. 6. Schlieren views of the thermocapillary vortex. a) During pulsating spreading. b) In the pseudo-uniform regime.

with the schlieren technique, is shown in Fig. 6b. The origin of this transition seems to be connected with the finite length of the channel, as it is possible to postpone the appearance of the homoclinic behavior by increasing this length. The conclusion is, therefore, that the finite size of the container affects propagation because the length of the eddy becomes comparable to the length of the channel. This propagation regime is named quasi-uniform in the literature because, in actual experiments, the flame undergoes a slow acceleration on approaching the end of the channel, in a stage which may span an appreciable fraction of the total spreading time.

4 Flow of the Liquid Underneath a Spreading Flame

To a large extent, the thermocapillary flow induced in the liquid fuel by the spreading flame can be analyzed separately of the more complex processes going on in the gas. Here, in order to uncouple the liquid and gas phase problems as much as possible, vaporization will be assumed to occur at a constant temperature T_v above the initial temperature of the liquid T_0, and the effect of the vaporization flux, which is obviously important in the dynamics of the gas, will be neglected in comparison with the larger fluxes prevailing in the bulk of the liquid. Neglecting, in addition, the effects of buoyancy and surface deformation, and using a reference frame moving with the vaporization front (where

the surface temperature first reaches the value T_v and vaporization begins), the nondimensional equations governing the motion of the liquid, assumed to be two-dimensional, are

$$\boldsymbol{\nabla} \cdot \boldsymbol{v} = 0, \tag{1}$$

$$\mathrm{Re}\left(\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} \boldsymbol{v}\right) = -\boldsymbol{\nabla} p + \nabla^2 \boldsymbol{v}, \tag{2}$$

$$\mathrm{Re}\,\mathrm{Pr}\left(\frac{\partial \theta}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} \theta\right) = \nabla^2 \theta \tag{3}$$

$$y = 0 : \begin{cases} v = 0, \quad \partial u/\partial y + \partial \theta/\partial x = 0, \\ \partial \theta/\partial y = q_g(x) \text{ for } x < 0 \quad \text{and} \quad \theta = 1 \text{ for } x > 0 \end{cases} \tag{4}$$

$$y = -1 : \quad u = U, \quad v = \theta = 0 \tag{5}$$

$$x \to -\infty : \quad u = U, \quad v = \theta = p = 0 \tag{6}$$

$$x \to \infty : \quad u = U, \quad v = 0, \quad \theta = 1 + y \tag{7}$$

where x and y are distances along and normal to the surface; $\theta = (T - T_0)/(T_v - T_0)$; and the variables $\boldsymbol{x}$, $\boldsymbol{v}$ and p are scaled with the factors h, $|\sigma'|(T_v - T_0)/\mu$ and $|\sigma'|(T_v - T_0)/h$, respectively, h being the depth of the liquid and $\sigma' = \mathrm{d}\sigma/\mathrm{d}T$. Finally, $q_g(x)$ is the heat flux reaching the nonvaporizing surface ($x < 0$), which should be determined from the analysis of the gas and is the only remaining link between the two phases. The two nondimensional parameters entering the problem are the Reynolds number $\mathrm{Re} = \rho|\sigma'|(T_v - T_0)h/\mu^2$ and the Prandtl number $\mathrm{Pr} = c\mu/k$, where ρ, μ, k and c are the liquid density, viscosity, thermal conductivity and specific heat. The nondimensional spread rate U is an eigenvalue to be determined with the condition that the solution be regular at the vaporization front $x = y = 0$, up to logarithms of $|\boldsymbol{x}|$. An analysis of the possible singularity of the solution at this point [15] shows that $\theta - 1 = O\left(|\boldsymbol{x}|^{1/2}\right)$ for very small values of $|\boldsymbol{x}|$, leading to a thermocapillary stress that diverges as $|\boldsymbol{x}|^{-1/2}$, for all but a special value of U.

Here, rather than deriving $q_g(x)$ from an analysis of the gas phase, a model distribution with constant shape and adjustable amplitude will be used, of the form $q_g(x) = q_0 \exp\left(-x^2/\sigma^2\right)$ where the width $\sigma \approx 0.1$ is chosen on the basis of order of magnitude estimates presented elsewhere [8]. Under stationary conditions, regularity at the origin will be achieved only for a certain $U = U(q_0; \mathrm{Re}, \mathrm{Pr})$.

A few stationary numerical solutions of (1)–(7) for $\mathrm{Re} = 5 \times 10^3$ and $\mathrm{Pr} = 10$ are displayed in Fig. 7. These solutions have been computed assigning a constant value to U and iterating on q_0 by means of a pseudo-transient method until a regular flow is obtained.

The asymptotic form of the solution of (1)–(7) for very large values of Re and moderately large Pr is of interest. The experimental and the numerical results show that the surface temperature falls rapidly to the left of the origin, so that the thermocapillary stress pushing the liquid toward the left imparts a total momentum flux $\int_{-\infty}^{0} \partial u/\partial y|_0 \, \mathrm{d}x = -1$ [from the second boundary condition (4)]. In the case of very small spread rates, the liquid recirculates with characteristic

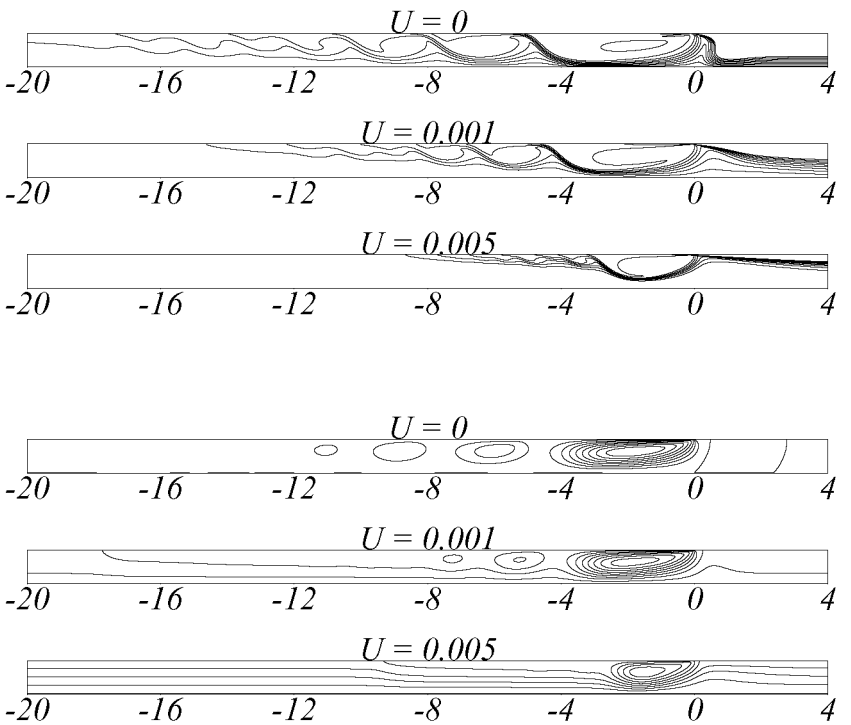

Fig. 7. Isotherms (three upper figures) and streamlines for the stationary flow with Re = 5000, Pr = 10 and three different values of the propagation speed. Only ten equispaced isotherms between $\theta = 0$ and $\theta = 0.3$ have been plotted in the upper figures.

velocity $u_0 = \mathrm{Re}^{-1/2}$ in a region of characteristic length $x_0 = \mathrm{Re}^{1/2}$. These estimates, which hold insofar as $U \ll u_0$, follow from the balance of convection and viscous forces in (2), $\mathrm{Re}\, u_0^2/x_0 = u_0$, and the condition $\mathrm{Re}\, u_0^2 = 1$, expressing that the overall momentum flux is still of the order of the thermocapillary force. This momentum will be ultimately lost because friction with the bottom gives rise to an overpressure that limits the horizontal extent of the flow to the left of the origin.

Nearer to the origin, the momentum given to the liquid is confined to a surface jet whose characteristic thickness, $\delta_j = \mathrm{Re}^{-1/3}(-x)^{2/3}$, and velocity, $u_j = \mathrm{Re}^{-1/3}(-x)^{-1/3}$, follow from order of magnitude balances similar to the ones above. If $U \gg u_0$, then the incoming flow (relative to the propagating front) reduces the size of the recirculation eddy to $x_0' = O\left(\mathrm{Re}^{-1}U^{-3}\right) \ll x_0$ [from the condition $u_j(x_0') = O(U)$], and the bottom ceases to play an important role when $U \gg \mathrm{Re}^{-1/3}$, because then $x_0' \ll 1$. In any case, the entrainment of the jet induces velocities of order $u_e = u_j\delta_j/(-x) = \mathrm{Re}^{-2/3}|x|^{-2/3}$ around the origin, for $|x|$ of order unity or small, and the condition $u_e = O(U)$ yields an estimate of

the extent of the recirculation eddy to the right of the origin: $x_1 = O(\mathrm{Re}^{-1}U^{-3/2})$ if $U \gg \mathrm{Re}^{-2/3}$, with $x_1 = O(1)$ otherwise.

Further estimates of the heat flux entering the liquid through the vaporizing surface between the origin and x_1, involving the energy equation (3), show that this heat flux reaches values of order $q_{max} = \mathrm{Pr}^{1/2}\mathrm{Re}^{2/3}\max(\mathrm{Re}^{1/3}U^{1/2}, 1)$ in a tiny region around the origin. Because of the regularity condition at the origin, q_{max} is also the order of the q_0 required for the front to propagate at the assumed velocity U. As can be seen, q_0 has a minimum value of order $\mathrm{Pr}^{3/2}\mathrm{Re}^{2/3}$ even when $U = 0$, and increases proportionally to $U^{1/2}$ when $U \gg \mathrm{Re}^{-2/3}$.

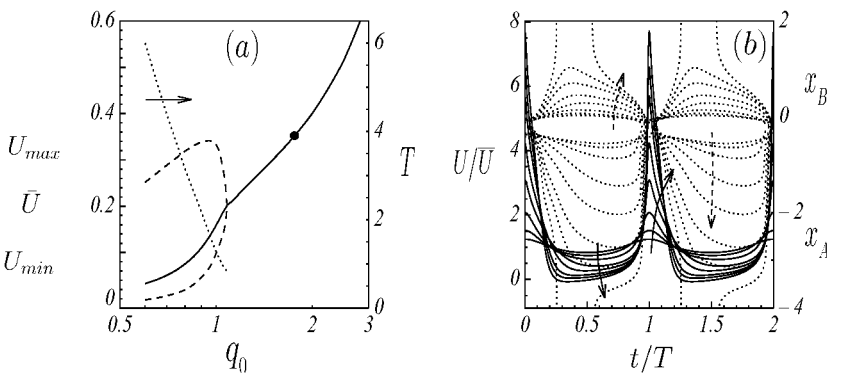

Fig. 8. (a) Nondimensional time average propagation speed ($\overline{U}$) as a function of q_0 for Re = 50 and Pr = 10 (solid); maximum and minimum speeds in the pulsating regime (dashed); and period of the oscillation (dotted). (b) Propagation speed, scaled with its time average value, as a function of time, scaled with the period, during two cycles of the oscillation for $q_0 = 1.07$, 1.05, 1. 0.9, 0.8, 0.7, 0.65 and 0.6, decreasing as indicated by the arrows (solid). The dotted curves give the distances from the front to the leading and trailing stagnation points.

Numerical solutions of (1)–(7) were also computed with a fully transient method for given values of q_0. The nondimensional propagation speed U obtained from these computations is plotted in Fig. 8a as a function of q_0 for Re = 50 and Pr = 10. Since q_g is nondimensionalized with $k(T_v - T_0)/h$, decreasing q_0 amounts to decreasing the liquid temperature T_0. The size of the recirculation bubble decreases with increasing q_0 until it disappears when $q_0 > q_{0b} \approx 1.76$ (black circle in fig. 7a). On the other hand, the flow becomes oscillatory when $q_0 < q_{0c} \approx 1.08$, the value of $\overline{U}$ in Fig. 8 being then the mean propagation speed, while the dashed curves give the minimum and maximum speeds in the cycle. The period T of the oscillation, nondimensionalized with $h/\overline{U}$, is given on the right side ordinate. The results in Fig. 8a are to be compared with the experimental data in Figs. 3 and 4, among others, though the present value of Re is substantially smaller than in the experiments.

The propagation speed scaled with $\overline{U}$ is plotted in Fig. 8b as a function of t/T during two periods of the oscillation. When q_0 is decreased sufficiently below q_{0c} the scaled speed displays the phases of rapid and slow (jump and crawling) propagation documented experimentally [3]. Also plotted in fig. 8b are the distances from the vaporization front to the leading and trailing stagnation points $x_{_A}(t/T)$ and $x_{_B}(t/T)$ (as they appear in the reference frame moving with the front), whose maxima increase with decreasing q_0 until they reach the boundaries of the computational domain (which extended from $x = -15$ to $x = 10$ in the present computations) during part of the cycle.

Evidently, the origin of these oscillations is to be sought in the dynamics of the liquid, because the gas plays a passive role in the present model, merely supplying a time independent heat flux for $x < 0$. Apparently the oscillations are connected with the existence of a small region of large spatial velocity variations around the vaporization front, and with the delay and feedback brought about by the recirculating flow.

5 Discussion of Results and Conclusions

The different regimes of flame spreading over liquid fuels have been studied experimentally and numerically. A series of transitions between these regimes has been identified as a cascade of bifurcations, indicating that a flame spreading in an elongated channel behaves as a dynamical system with only one effective spatial dimension directed along the propagation. The physical constraints in the transverse directions impede the development of instabilities along them. The initial temperature of the condensed fuel is a convenient control parameter whose variation allows the system to display its different spreading regimes. Starting at high values of the fuel temperature and decreasing it, the first observed transition marks the beginning of the truly spreading regimes: from a pure gas phase propagation to a spreading regime where the condensed fuel plays an active role by assisting the flame propagation. The next transition is a steady bifurcation to a thermocapillary regime, and indicates that thermocapillary stress is the driving mechanism at these temperatures, where it leads to the formation of a characteristic eddy structure. This transition occurs when a characteristic thermocapillary velocity, defined on the basis of the surface tension gradient, becomes comparable to the flame spreading velocity. The ratio of the two velocities takes the same value at the transition point, up to the experimental error, for the four alcohols investigated. This result suggests the existence of a universal critical value of this ratio indicating the appearance of the thermocapillary eddy. As remarked time ago [2], this second transition happens at a temperature close to the flash point temperature. The coincidence is surely not accidental, and the reason may be the opposite effects that transport of heat by thermocapillary convection has on the process of flame propagation and on the process of flame initiation. In the first case, convection preheats the fuel with some of the heat released in the combustion region, and thus enhances propagation. In the second case, the thermocapillary currents carry heat away from the initia-

tion region, opposing the progress of the reaction. The evidence of some of these features has received support from the numerical computations. In spite of the strong simplifications introduced in the model regarding the coupling with the gas phase, the numerical results do not only account for the steady bifurcation to the thermocapillary regime but also are able to predict the Hopf bifurcation to the pulsating regime and the evolution towards relaxation oscillations. All of this points out that the physical reasons behind the appearance of the pulsating regime are related to the instability of the dynamic heat balance in the liquid phase, with the gas phase playing a slave role. This conclusion contrasts with the common belief that the oscillation is driven alternatively by the gas phase and the liquid phase [3]. Many details of this transition, as well as aspects of the homoclinic bifurcation, are in need of further investigation and will be the subject of follow on studies.

This work was supported by DGES grant PB98-0142-C04.

References

1. A. Liñán: In *Combustion in High Speed Flows, proc. ICASE Meeting, 12–14 Oct. 1992*, ed. by J. Buckmaster, R. Kumar (Kluwer 1994) pp. 461–476.
2. F. A. Williams: *Combustion Theory*, 2nd edn. (Benjamin-Cummings, Menlo Park, Ca 1985).
3. H. D. Ross: Prog. Energy Combust. Sci. **20**, 17 (1994).
4. W. A. Sirignano, I. Glassman: Combust. Sci. Technol. **1**, 307 (1970).
5. K. Akita: Proc. 14th Symp. on Combustion, 1075 (1973).
6. P. L. Garcia-Ybarra, J. C. Antoranz, V. Sankovitch, J. L. Castillo: Phys. Rev. E **49**, 5225 (1994).
7. P. L. Garcia-Ybarra, J. L. Castillo, J. C. Antoranz, V. Sankovitch, J. San Martin: Proc. 26th Symp. on Combustion, 1469 (1996).
8. F. J. Higuera, P. L. Garcia-Ybarra: Combust. Theory Modelling **2**, 43 (1998).
9. E. Degroote, P. L. Garcia-Ybarra: Eur. Phys. J. B **13**, 381 (2000).
10. E. Degroote, P. L. Garcia-Ybarra: Phys. Fluids (submitted, 2000).
11. J. N. deRis: Proc. 12th Symp. on Combustion, 241 (1969).
12. F. J. Higuera: Combust. Theory Modelling **3**, 147 (1999).

Applications of Mathematical Modelling to Biological Pattern Formation

Philip K. Maini

Centre for Mathematical Biology, Mathematical Institute, Oxford, OX1 3LB, UK

Abstract. The formation of spatiotemporal patterning in biology has intrigued experimentalists and theoreticians for many generations. Here we present a brief review of some mathematical models for pattern formation and then focus on three models which use the phenomenon of chemotaxis to generate pattern.

1 Introduction

One of the characteristic features of nature is the enormous diversity of spatial structure and form present in plants and animals. Understanding how such patterns arise is the central goal of a large body of experimental and theoretical research. Although the spectacular advances over the past decade in molecular biology have resulted in detailed descriptions of the spatiotemporal dynamics of various genes and their products, such studies shed only very limited light on the underlying origin of these dynamics. The latter is the interest of the mathematical biologist, and many models have been proposed to describe the formation of biological patterns.

One of the most famous examples of spatiotemporal pattern formation is the Belousov-Zhabotinsky chemical reaction, in which bromate ions oxidise malonic acidic in a reaction catalysed by iron which, depending on its state, can assume two different colours – reddish-orange or blue. By observing the colour changes, one can see that this reaction exhibits a wide range of spatiotemporal patterning, such as propagating fronts, spiral waves, target patterns and toroidal scrolls. This reaction may be modelled mathematically by a coupled system of nonlinear partial differential equations which account for the reaction kinetics between species and also their diffusion. Solutions of the resultant system of equations exhibit the experimentally observed phenomena for chemically realistic parameter values (for review, see [1,2] and references therein).

Such oscillatory and wave-like patterns also arise in physiology. For example, the heart beats in response to an electrical stimulus which moves in a wave-like fashion across the heart. This may be modelled phenomenologically by a system of equations which are not dissimiliar to those mentioned above, and the resultant models have been used to try to understand abnormal heart behaviour, such as fibrillation, which can lead to death (for review, see [3]). Spiral waves also play a key part in the life cycle of the slime mould *Dictyostelium discoideum* and we will focus on this in section 4.

The above are examples of patterns that change in space and in time. Perhaps the most elegant model for temporally stationary spatial patterns is that

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 205–217, 2001.
© Springer-Verlag Berlin Heidelberg 2001

proposed by Turing [4]. He showed that a pair of reacting and diffusing chemicals could evolve from initial nearly uniform steady state chemical concentration profiles to spatially varying profiles. He hypothesized that if one of the reacting pair of chemicals was a growth hormone, then this underlying spatially heterogeneous chemical profile would result in spatially non-uniform growth leading to the development of structure, or a spatial pattern. Hence he termed these chemicals *morphogens*.

Turing's analysis showed that it was possible to chose a reaction-diffusion system in which the kinetics yielded a steady state which was stable in the absence of diffusion. However, the introduction of diffusion, which we intuitively think of as a stablizing process, causes instability. Hence, this was termed *diffusion-driven instability* and is an example of self-organisation, or an emergent property.

Since Turing's seminal paper there has been an enormous amount of literature in which different types of nonlinear models are derived and analysed, all based on his original idea. Although the existence of morphogens still remains a controversial issue, there are a number of candidates (see, for example, [5]). Some 40 years after Turing predicted the phenomenon of diffusion-driven instability, it was observed in a chemical system, the Chloride-Iodide-Malonic Acid (or CIMA) reaction [6,7].

In this paper we present a brief review of previous models that have been proposed to account for spatial patterning in biology. We then consider, in a little more detail, three different models which have the common underlying theme of involving chemotaxis, a widely-used mechanism for directed cell movement which has been extensively studied from both the experimental and theoretical viewpoints [8,9].

2 Models for Spatial Pattern Formation: A Brief Review

A number of models have been proposed to account for spatial pattern formation and regeneration in several areas of developmental biology. This is now a huge area of research and to aim to review all such models in one short paper is too ambitious. Therefore, here we present only a few of the models in order to give the reader a flavour of the area.

Broadly speaking, most models may be classified as either *chemical prepattern* models or *cell movement* models. We consider each in turn:
(i) Chemical pre-pattern models: These models assume that a spatial pattern in some chemical (termed a morphogen) is set up and cells respond to this pattern by differentiating accordingly, that is, cells are distributed in a spatially uniform manner, and the spatial heterogeneity arises due to the information (termed *positional information*) cells receive from the chemical pre-pattern. There are two main ways in which such a spatially non-uniform pattern in chemical concentration may arise: It may be due to a simple source-sink mechanism, coupled with diffusion and degradation, which cells then interpret via a complex mechanism involving multiple thresholds [10], or it may arise due to a reaction-diffusion, or Turing system. Such a pattern, in its simplest one-dimensional form, consists

of a series of peaks and troughs, which require only one threshold of cellular interpretation in order to lead to a spatial pattern of cell differentiation. These models may be couched in terms of partial differential equations, or discretized and analysed as cellular automata [11].
(ii) Cell movement models: These models assume that a spatial pattern arises in cell density, and cells then differentiate in a density-dependent manner; for example, those cells in high density aggregates differentiate while cells in low density areas remain undifferentiated. There are a number of ways in which cell aggregation can occur. Cells can move in response to mechanical and chemical cues, such as gradients in certain chemical concentrations and/or adhesive sites in the extracellular material in which they move. A number of authors have studied these types of models and shown that they can lead to complex spatial patterns in cell density (see, for example, [12,13,14]).

Models formulated on the hypotheses that cells move to minimize energy have also been shown to give rise to pattern formation and cell sorting behaviour consistent with a number of experimental observations [15,16,17,18].

Chemical pre-pattern and cell movement models have been used to account for a wide range of pattern formation and morphogenetic phenomena, including gastrulation, pattern formation and regulation in Hydra, skeletal patterning in the limb, animal coat markings, feather formation on dorsal chick skin, tooth primordia formation in alligators, to name but a few (see [19], and references therein, for further details).

The above models are based on very different biological hypotheses and it is still an area of controversy as to which is correct. Although it appears that in alligators, salamanders and certain fishes, the pigmentation patterns are due to cell movement, in many other cases a pattern of cell condensations and of certain chemicals are observed. However, as the time course of these patterning events has not yet been resolved (due to technical difficulties) it is still not clear what is cause and what is effect [20].

Although the biology underlying these models are different, several of them are based on *short-range activation, long-range inhibition*, which is a general mechanism for generating spatial pattern. Mathematically, the reaction-diffusion models and several of the cell movement models are coupled systems of nonlinear partial differential equations. Due to their complexity, most mathematical analyses have thus far consisted of determining linear stability and primary bifurcation points. One finds that for both model types, the patterns produced in the vicinity of a primary bifurcation point are eigenfunctions of the Laplacian satisfying the appropriate boundary conditions. This can be confirmed by numerical simulation. Therefore, both model classes make similar patterning predictions. Although this means that it is difficult to use these types of models to distinguish between hypotheses, it does make it possible to make predictions that are independent of the underlying biological hypotheses. This has been exploited to derive certain *developmental constraints* [21]. Perhaps the best known of these is that spotted animals with striped tails are much more common than striped animals with spotted tails [2].

The above models consider pattern formation at a macroscopic level and therefore cannot account for patterning on a finer scale. In many developing tissues, adjacent cells diverge in character to create a fine-grained pattern of cells in contrasting states of differentiation. For example, in the developing nervous system, nascent neural cells inhibit their neighbours from becoming committed to a neural fate, resulting in a fine-grained differentiation pattern. This is an example of the general lateral inhibition principle that underlies many such patterns. The appropriate type of mathematical model in this case is a coupled system of ordinary differential equations characterising the temporal changes of key chemicals or proteins in each individual cell, with cells being coupled by nearest neighbour interactions [22].

3 Primitive Streak Formation

Many pattern formation and morphogenetic studies focus on aspects of avian development. Although there are detailed experimental descriptions of the events occurring in the early stages of chick development, less is known about the mechanisms that underly these events. One of the earliest pattern forming events, which also sets the stage for the proper coordination of subsequent development, is the formation of the primitive streak. During streak formation, a group of specialized cells at the posterior margin of the disc-like blastoderm moves anteriorly across the disc to about three-fifths the way across, and then it regresses. During regression, large scale cell movements occur turning the two-dimensional embryo into a three-dimensional structure. The movement of the streak is an example of a novel type of travelling wave. Travelling waves of constant speed and profile are well-known to occur in scalar reaction-diffusion equations and they have been applied to account for the propagation of a favoured gene through a population, while such waves in coupled systems have been used to account for the invasion of one species by another [2], or the closure of wounds [23].

To date, little in known about the mechanisms controlling progression and regression in the primitive streak. Recently, a simple model based on cell-chemotaxis was proposed and analysed, and shown to exhibit many of the features observed experimentally. Here, we briefly review the main points of the model and refer the reader to the original paper for full details [24].

The model postulates that the density $n(\mathbf{x}, \mathbf{t})$ at spatial position $\mathbf{x}$ and time t of the subpopulation of cells that comprise the primitive streak secrete and respond to a chemoattractant which has concentration $u(\mathbf{x}, \mathbf{t})$ and takes the general form:

$$\frac{\partial n}{\partial t} = D_n \nabla \cdot (\nabla n - n\chi(u)\nabla u) + f_1(n, u) \tag{1}$$

$$\frac{\partial u}{\partial t} = D_u \nabla^2 u + f_2(n, u) \tag{2}$$

where D_n and D_u are constant diffusion coefficients and the second term on the right-hand side of the first equation models movement up the gradient of the

chemical (chemoattraction) with chemotactic sensitivity. Modelling how cells respond to external signalling cues is a very active area of research at the moment. To date, most models are phenomenological, although recently it has been shown how rules of motion at the individual cell level may be translated, using the ideas of biased random walks, to the cell population level [25]. For simplicity, here we assume that the sensitivity is a constant χ_0. Net cell and chemical production are modelled by the terms f_1 and f_2 respectively. It is known that cell motion during these early stages of development occurs on a faster timescale than cell division, so f_1 can be taken to be 0. The chemical kinetics is chosen to be of the form:

$$f_2(n,u) = g(n) - \beta nu - \delta u \tag{3}$$

where $g(n)$ accounts for secretion of the chemical by cells, and is assumed to depend sigmoidally on n. (In a very simple case, it can be taken to be 0 for $n < n_1$, and a non-zero constant γ for $n \geq n_1$ for some constant threshold cell density level n_1). The other two terms in the equation above model chemical degradation due to the cells and to natural decay.

The appropriate boundary conditions are zero flux. If the model is considered on a one-dimensional spatial domain, then appropriate initial conditions for n are that $n(x,0) = n_i$ for $x < x_i$ but zero otherwise. If the initial chemical concentration is taken to be spatially uniform and the model equations solved numerically, it is found that this simple model exhibits progression followed by regression. The mechanism underlying this behaviour is that the initial cell density at the origin degrades the chemical concentration creating a gradient, up which cells start to move. At some later stage, however, the chemical concentration ahead of the advancing cell wave drops to a value below that behind the wave, reversing the chemical gradient and hence cell motion. It is possible to determine the parameter space in which this behaviour will occur and it turns out to be a robust phenomenon. The time of reversal can be approximated in terms of the parameters of the model as can the speed.

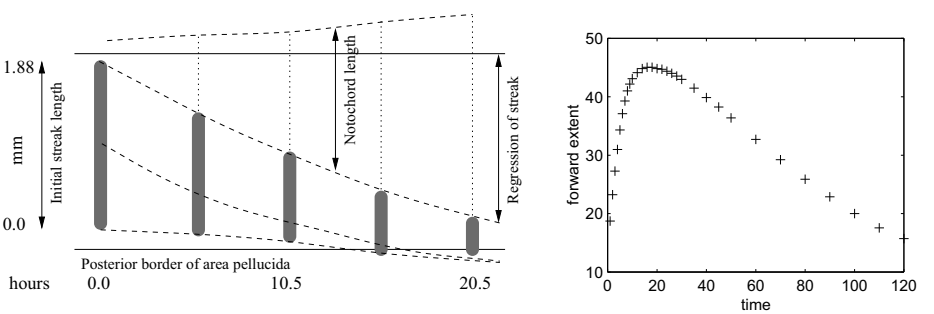

Fig. 1. Comparison of experimental data with numerical simulations of model equations (1)-(2) in one dimension showing the spatial extent of primitive streak regression. Reproduced from [24] with permission.

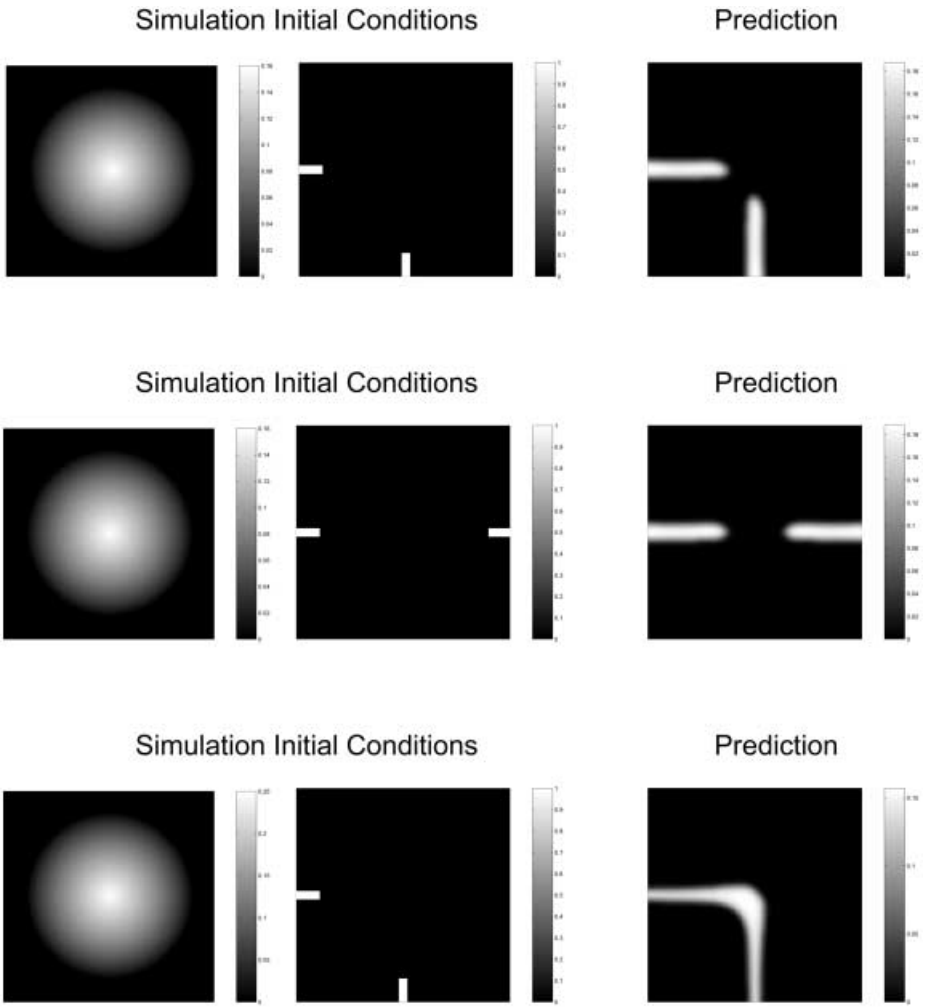

Fig. 2. Numerical simulations of model equations (1)-(2) for the case where a donor primitive streak is implanted at different places in a host. The host primitive streak is at the left-hand margin of the domain. The resulting "prediction" agrees with experimental observations. Note that two of the experiments give different predictions: the top panel predicts two ingrowing streaks, the bottom panel predicts fused streaks. The outcome of the numerical simulations is crucially dependent on the initial conditions and reflects the experimental observations that the outcome is either two separate streaks or a fused pair. The left-most panels show the concentration of chemoattractant. Reproduced from [24] with permission.

Of course, on a two-dimensional spatial domain, the above initial conditions result in a spreading disc of cells, whereas the primitive streak itself is a tightly compacted column of cells. By assuming that the initial chemical concentration

on a two-dimensional domain has a maximum at the centre of the domain, however, numerical simulations show that the model exhibits behaviour in close agreement with experimental observations (Figs 1 and 2). Fig. 3 shows how the model can be used to make experimentally testable predictions.

Hence, this simple model shows that the hypothesis that primitive streak formation is due to chemotaxis is a plausible one, and it makes experimentally testable predictions.

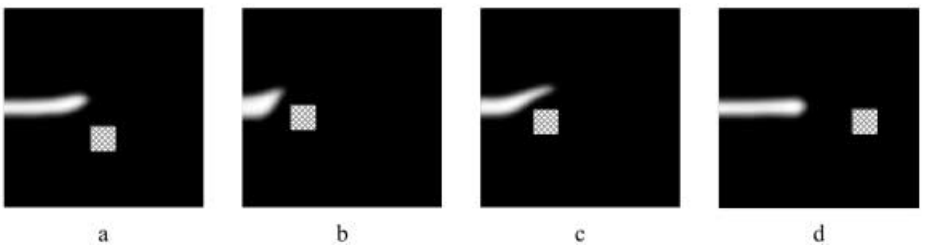

Fig. 3. Model predictions for streak development after experimental ablation (hatched area) of certain portions of the blastodisc. Reproduced from [24] with permission.

4 Cell Aggregation in Slime Mould

The amoebae of the cellular slime mould *Dictyostelium discoideum* (Dd) have a remarkable life cycle. In response to starvation conditions, the amoebae secrete a chemical, cyclic 3'5'-adenosine monophosphate (cAMP), which diffuses through the medium in the form of spiral waves. This is a chemottractant for the cells and causes the phenomenon of cell streaming, leading to cell aggregation and finally to the formation of a multicellular organism, known as the slug, composed typically of $10^4 - 10^5$ cells. In the body of the crawling slug, the cells differentiate into pre-spore and pre-stalk cells. Finally, the slug culminates into a fruiting body in which the stalk cells lift the spore cells, aiding their dispersal. The amoebae can remain in the spore stage for many years before emerging under favourable conditions.

This developmental program has been studied in great detail by experimentalists and theoreticians alike because it contains several elements, for example, signal propagation and tranduction, cell movement, sorting and differentiation, that are common to higher organisms. Hence it serves as an excellent model paradigm.

We briefly summarise the modelling here and refer the reader to [26,27,28,29] and references therein for full details. The molecular mechanisms of the cAMP signalling dynamics have been studied in detail and a number of models have been proposed. The simplest of these exploit the different time scales on which reactions occur to use a quasi-steady state assumption to reduce the order of

the system, leading to a coupled pair of ordinary differential equations in which the variables are extracellular cAMP concentration and density of cAMP receptors. This is an excitable system and exhibits propagating spiral waves, in good agreement with those observed experimentally.

To model cell streaming, one must couple cell movement and chemotaxis to this excitable system. This may be done by considering the cells as discrete entities for which cAMP-dependent movement rules are specified, and the cAMP concentration profiles are determined by a numerically computed finite-difference approximation to the above excitable system. This method has the advantage of allowing one to follow individual cells and compare at a detailed level, the model with experimental observations. However, the form of the model is such that little mathematical insight can be gained. Alternatively, the cells can be modelled as a continuum. This has the disadvantage of allowing less detailed comparison with experiment, but the resultant system of coupled partial differential equations can be analysed mathematically to some extent [28]. We refer the reader to [29] for more details of these two alternative methods and a comparison between them.

Here, we consider the following continuum model [26,27]:

$$\frac{\partial n}{\partial t} = \nabla \cdot (\mu \nabla n - \chi(v) n \nabla u) \tag{4}$$

$$\frac{\partial u}{\partial t} = \lambda[\phi(n) f_1(u,v) - (\phi(n) + \delta) f_2(u)] + \nabla^2 u \tag{5}$$

$$\frac{\partial v}{\partial t} = -g_1(u)v + g_2(u)(1-v), \tag{6}$$

where n, u and v denote cell density, extracellular cAMP concentration and fraction of active cAMP receptors, respectively. The first equation is the cell conservation equation and incorporates diffusion, with cell diffusion coefficient μ, and chemotaxis, with chemotactic sensitivity denoted by $\chi(v)$. The simplest form of $\chi(v)$, namely a constant, is not appropriate here, because it would mean that cells would move both in response to the front of the signalling cAMP wave and also in the waveback. This contradicts experimental observations, which show movement only in the wavefront. Moreover, it predicts that the net cell motion will be in the direction the signal is travelling, rather than in the direction from where the signal originated. This was known as the "chemotactic paradox" and can be resolved by accounting for adaptation. Hence, the chemotactic sensitivity is taken to be of the form $\chi(v) = \chi_0 v^m/(A^m + v^m)$, $m > 1$, where χ_0 and A are positive constants. This functional form reflects the hypothesis that an appreciable chemotactic response requires a minimal fraction of active receptors, yet, for a large fraction of active receptors, the response saturates.

The second and third equations are a simplified version of the excitable model of the cAMP-cell receptor dynamics [30] modified by cell density effects. The first term on the right-hand side of the second equation is the cAMP production term and assumes that the rate of production per unit cell density is $f_1(u,v)$, where $f_1(u,v) = (bv + v^2)(a + u^2)/(1 + u^2)$. This models autocatalytic cAMP production with saturation, mediated by cAMP binding to active receptors. The rate of cAMP degradation per cell is taken to be $f_2(u) = du$. The cell density

dependence is taken to be $\phi(n) = n/(1 - \rho n/(K + n))$, while δ accounts for cAMP degradation in the absence of cells. The parameters a, b, d, ρ, K, δ and λ are all positive constants.

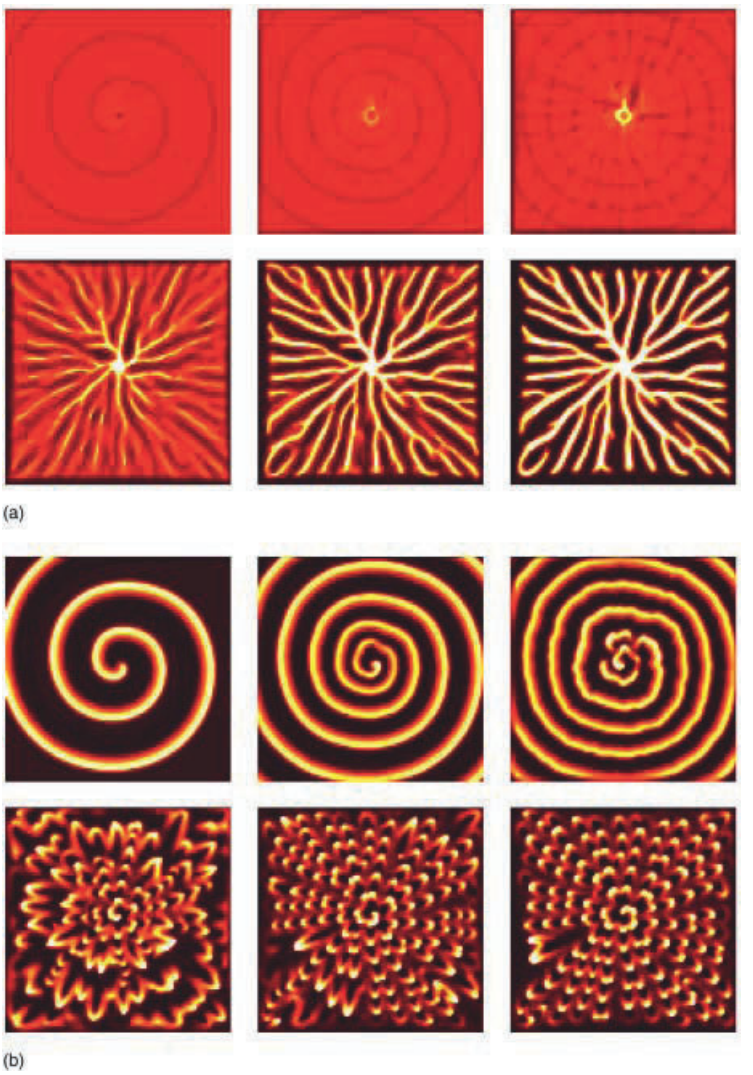

Fig. 4. Typical numerical simulation of model equations (4)-(6) showing the time evolution (left to right) of (**a**) cell density and (**b**) cAMP. Note the phenomenon of cell streaming (panels 5 and 6 in (**a**)).

The first term on the right-hand side of the third equation accounts for receptor densitization and, assuming the law of mass action, $g_1(u) = k_1 u$. The second

term models resensitization of the desensitized $(1 - v)$ fraction of receptors, at the constant rate $g_2(u) = k_2$. The parameters k_1 and k_2 are positive constants.

Good estimates of most of the parameters are available from the experimental literature and we refer the reader to the original paper [26] for these estimates and also for full details on the derivation of each term, which we have only sketched above. Substituting these parameter values into the model, we find that the model captures the key features of the cell streaming process. Fig. 4 shows a typical model simulation.

It is found that the initial uniform state is unstable from the outset, suggesting that the coupled dynamics of cAMP wave propagation and cell movement exhibit a patterning instability perpendicular to the direction of wave propagation. This has been investigated in an analytically-tractable caricature of the above model from which it is possible to derive explicit conditions on the parameters for the uniform steady state to go unstable, and to predict the wavelength of the fastest growing mode from the dispersion relation. The results are in good agreement with the numerical simulations of the full model above and experimental observations [27].

Numerical simulations show that low initial cell densities lead to the formation of a central hole agreeing with experimental observations in which the excitability of the medium is lower. It can be shown mathematically that, in the model, this is equivalent to lowering the cell density.

The model is not only consistent with a number of experimental observations, it can also provide alternative explanations for some observed phenomena. For example, it is known that the wavespeed and wavelength of the spiral patterns decrease as cell streaming procedes. This has been explained by assuming that complex biochemical changes must be occurring in the medium resulting in the parameters changing their values. However, the above model, in which the parameter values are kept constant, exhibits identical behaviour. It can be shown mathematically that this occurs due to the spiral dynamics.

5 Pigmentation Patterning in Fishes

A number of models have been proposed to account for the spectacular variety of pigmentation patterns seen on animals. These models show that the complexity of the pattern depends on the size of the domain at the instance when the pattern is laid down. None of these studies considered domain growth during the pattern formation process as it was assumed that, particularly in mammals, domain growth simply caused the pattern to change quantitatively, not qualitatively. However, it was shown recently that the pigmentation patterns on certain fishes change qualitatively with domain growth [31]. The authors studied the striped patterns on the marine angelfish, *Pomacanthus*, and observed that as the fish grows, its stripes grow wider apart until they are about twice the distance apart as in the juvenile case. At this stage, new stripes appear between the existing stripes in such a way that the original interstripe spacing is restored. The new stripes are thinner than the pre-existing stripes but they gradually broaden.

It was shown in [31] that this mode-doubling phenomenon was consistent with the Turing theory. The general Turing model takes the form:

$$\frac{\partial u}{\partial t} = D_u \nabla^2 u + f(u, v) \tag{7}$$

$$\frac{\partial v}{\partial t} = D_v \nabla^2 v + g(u, v) \tag{8}$$

where f and g are the reaction kinetics for chemicals with concentrations u and v. The authors considered very simple piecewise linear kinetics and included domain growth in an *ad hoc* way in their simulations. Their results showed that as the domain grew, new stripes were inserted in a fashion similar to that observed on the fish, except that the inserted stripes had exactly the same width as the pre-existing stripes. This is to be expected as the wavelength of the pattern in a Turing system depends only on the model parameters.

It is not yet clear if the pigmentation pattern in *Pomacanthus* arises due to a chemical pre-pattern (as hypothesized in this model) or due to cell movement. Observations in other fish, such as zebrafish, suggest that cell movement may play an important role in pigmentation patterning. Therefore, a model was developed in [32] which accounted for cell movement and modelled domain growth in a mathematically rigorous way. It was shown that, under domain growth, the concentration c of a chemical in a reaction-diffusion system satisfies an equation of the form:

$$\frac{\partial c}{\partial t} + \nabla \cdot (\underline{a} c) = D \nabla^2 c + f(\underline{u}) \tag{9}$$

where D is the diffusion coefficient, $\underline{a}$ is the velocity of growth of the domain, and $\underline{u}$ is the vector of chemical concentrations. It can be shown that this system undergoes mode doubling for quite general forms of reaction kinetics and that this phenomenon can arise not only via insertion, but also by splitting [33].

Cell movement may be incorporated into this model by assuming that cell density, n, satisfies an equation of the form:

$$\frac{\partial n}{\partial t} + \nabla.(n\underline{a}) = D_n \nabla^2 n - \nabla.[\chi_u(u) n \nabla u + \chi_v(v) n \nabla v] + n(r_x + r_y) \tag{10}$$

where $\underline{a} = (r_x x \mathbf{i} + r_y y \mathbf{j})$ is the velocity field (assumed uniform for simplicity) generated by growth, r_x and r_y are growth rates in the x and y directions, respectively, u and v are the concentrations of the chemicals in the reaction-diffusion system, $\chi_u(u)$ and $\chi_v(v)$ are the chemotactic sensitivities. The last term on the right-hand side is a cell source term chosen to ensure conservation of cell density.

In [32] it is shown that not only can a model of this form produce mode doubling behaviour (obviously possible since the underlying Turing model is unchanged) but that, for biologically realistic parameters, the inserted stripes are narrower to begin with and then gradually broaden. The reason behind this behaviour is that although the newly inserted stripes of chemical concentration in the underlying reaction-diffusion system have width identical to the pre-existing

stripes, the cell rearrangement induced by this new pattern occurs slowly, for realistic diffusion coefficients. As a result, the new inserted stripes in cell density are thinner than the existing stripes, broadening only gradually (Fig. 5). In fact, it can be shown that for suitable choice of chemotactic sensitivites $\chi_u(u)$ and $\chi_v(v)$, this model can give rise to cell density steady state patterns of thick and thin stripes [34].

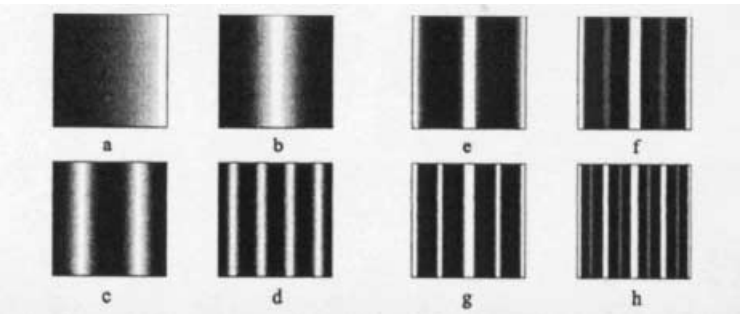

Fig. 5. Cell density plots in a numerical simulation of the fish pigmentation model (equation (10) plus corresponding Turing system) showing stripe insertion as the domain grows (**a**) - (**h**). For convenience the domain has been rescaled to a fixed size. Reproduced from [34] with permission.

6 Discussion

The models presented in sections 3-5 share the common underlying process of chemotaxis but give rise to very different patterns due to the different dynamics that generate the pattern in chemoattractant. In the model for primitive streak a simple gradient in chemoattractant concentration is set up via degradation and diffusion. In the case of the slime mould, spiral waves of cAMP concentration are the result of the cAMP-active cAMP receptor system being an excitable medium, while in the fish pigmentation model, the chemical patterns are set up via diffusion-driven instability. The latter model also incorporated domain growth. One of the main future challenges in this general area is the detailed study of growth. In the model presented, growth was prescribed, for simplicity. In actual fact, of course, growth will result from the internal dynamics and this presents a new modelling challenge [35].

It is clear that biology offers mathematics a wide variety of novel and challenging problems. It is natural to ask what mathematics offers biology. Most biological processes interact in a nonlinear way. Therefore, to forecast the consequences of these interactions one must use a language in which nonlinearities can be manipulated. At the moment, mathematics is the only language that can do so. Therefore mathematics can be used to test hypothesis and also to make

experimentally testable predictions. The three models presented in sections 3-5 illustrate how mathematics may be used in this way.

References

1. B.R. Johnson, S.K. Scott: Chemical Soc. Reviews, **25**, 265 (1996)
2. J.D. Murray: *Mathematical Biology*, 2nd edn. (Springer-Verlag 1993)
3. A.V. Panfilov, A.V. Holden (eds): *Computational Biology of the Heart*, (John Wiley & Sons, Chichester 1997)
4. A.M. Turing: Phil. Trans. Roy. Soc. Lond. B **327**, 37-72 (1952)
5. L.G. Harrison: *Kinetic Theory of Pattern Formation*, (Cambridge University Press, 1993)
6. V. Castets, E. Dulos, J. Boissonade, P. De Kepper: Phys. Rev. Lett., **64**, 2953 (1990)
7. Q. Ouyang, H. L. Swinney: Nature (London) **352**, 610 (1991)
8. J. P. Armitage, J.M. Lackie: *Biology of the Chemotactic Response*, (Cambridge University Press 1990)
9. W. Alt, A. Deutsch, G. Dunn (eds): *Dynamics of Cell and Tissue Motion*, (Birkhäuser 1997)
10. L. Wolpert: J. theor. Biol., **25**, 1 (1969)
11. J.B.L. Bard: J. theor. Biol., **93**, 363 (1981)
12. E.F. Keller, L.A. Segel: J. theor. Biol., **30**, 235 (1971)
13. P.K. Maini, M.R. Myerscough, K.H. Winters, J.D.Murray: Bull. Math. Biol., **53**, 701 (1991)
14. G.F. Oster, J.D. Murray, A.K. Harris: J. Embryol.exp. Morph., **78**, 83 (1983)
15. M.S. Steinberg: J. exp. Zool., **173**, 395 (1970)
16. D. Sulsky, S. Childress, J.K. Percus: J. theor. Biol., **106**, 275 (1984)
17. G. Cocho, R. Pérez-Pascual, J.L. Rius: J. theor. Biol., **125**, 419 (1987)
18. G. Cocho, R. Pérez-Pascual, J.L. Rius, F. Soto: J. theor. Biol., **125**, 437 (1987)
19. P.K. Maini: 'Mathematical models for biological pattern formation'. In: *On Growth and Form: Spatio-temporal Patterning in Biology.* ed. by M.A.J. Chaplain, G.D. Singh, J.C. McLachlan (John Wiley & Sons, Chichester 1999) pp. 111-128
20. P.K. Maini, M. Solursh: Int. Rev. Cytology, **129**, 91 (1991)
21. G.F. Oster, J.D. Murray: Zool., **251**, 186 (1989)
22. J.R. Collier, N.A.M. Monk, P.K. Maini, J.H. Lewis: J. theor. Biol., **183**, 429 (1996)
23. P.D. Dale, P.K. Maini, J.A. Sherratt: Math. Biosciences **124**, 127 (1994)
24. K.J. Painter, P.K. Maini, H.G. Othmer: Bull. Math. Biol. **62**, 501 (2000)
25. H.G. Othmer, A. Stevens: SIAM J. Appl. Math. **57**, 1044 (1997)
26. T. Höfer, J.A. Sherratt, P.K. Maini: Proc. Roy. Soc. Lond. B **259**, 249 (1995)
27. T. Höfer, J.A. Sherratt, P.K. Maini: Physica D **85**, 425 (1995)
28. T. Höfer, P.K. Maini: Phys. Rev. E **56**, 2074 (1997)
29. J.A. Sherratt, J.C. Dallon, T. Höfer, P.K. Maini: 'Mathematical modelling of signalling in Dictyostelium discoideum'. In *Microbial Signalling and Communication.* ed. by R. England, G. Hobbs, N. Bainton, D. McL. Roberts (Cambridge University Press 1999) pp. 241-254
30. J.-L. Martiel, A. Goldbeter: Biohys. J. **52**, 807 (1987)
31. S. Kondo, R. Asai: Nature **376**, 765 (1995)
32. K.J. Painter, P.K. Maini, H.G. Othmer: Proc. Natl. Acad. Sci. **96**, 5549 (1999)
33. E.J. Crampin, E.A. Gaffney, P.K. Maini: Bull. Math. Biol. **61**, 1093 (1999)
34. K.J. Painter, P.K. Maini, H.G. Othmer: J. Math. Biol. (accepted)
35. R. Dillon, H.G. Othmer: J. theor. Biol. **197**, 295 (1999)

The Artistry of Bacterial Colonies and the Antibiotic Crisis

Ido Golding and Eshel Ben-Jacob

School of Physics and Astronomy, Raymond & Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University

1 Introduction

Since the beginning of massive usage of antibiotics during World War II we have witnessed a dramatic evolutionary event – the emergence of multiple drug resistant bacteria. The bacteria are capable of developing antibiotic resistance at a higher rate than scientists develop new drugs [1, and references therein. See also the UN's World Health Report 1996]. We seem to be loosing a crucial battle on our health. To reverse this course of events, we have to "outsmart" the bacteria by taking new avenues of study which will lead to the development of novel strategies to fight them.

1944 was the first year in which an antibiotic, penicillin, was produced in large enough quantities to be widely available. Soon after other antibiotics became commercially available and medicine had entered a golden era, as it became possible to cure deadly bacterial diseases that had plagued mankind throughout history. Along with the medical use came misuse; antibiotics are often prescribed with no medical justification, as a blanket therapy or as a preventive measure; they are given to animals to promote growth rather than treat disease; and they have other uses, e.g. the drugs are sprayed onto fruit trees to control bacterial infections [2,3].

The result of this continuous antibiotic pressure on the bacteria is the increasing number of strains of disease-causing bacteria that can resist multiple drugs. The bacteria have developed their resistance in antibiotic-filled human bodies, in animals and in the fields. Now we are experiencing an alarming emergence of common but no longer curable bacterial infections. The bacteria do not develop drug-resistance genes by Darwinian natural selection (gradual accumulation of random mutations) and do not compete for survival in the face of antibiotic attack. Rather, they use extensive chemical and genetic communication to redesign genes, pack them in an "off the shelf" resistance cassettes ('plasmids') and distribute them among peers [4,3]. Often one bacterium will pass resistance traits to others by giving them a useful plasmid. Viruses occasionally transfer resistance genes by extracting a gene from one bacterial cell and inject it into a different one. In addition, after a bacterium dies and releases its contents into the environment, another bacterium will occasionally take up a liberated gene for itself. All these methods of transferring genetic information enable bacteria to communicate not only with bacteria of the same species, but also with rela-

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 218–241, 2001.
© Springer-Verlag Berlin Heidelberg 2001

tive and foreign species, creating a "World Wide Web" of genetic communication [5,6,4].

The ways bacteria deal with antibiotic is as complex as the ways antibiotic harm bacteria. Antibiotics kill bacteria or inhibit their growth in several ways [7]: 1) by interfering with cell wall synthesis 2) by reducing plasma membrane integrity 3) by inhibiting nucleic acids synthesis 4) by interfering with ribosomal function 5) and by inhibiting folate synthesis (folate molecule is a carrier of carbon in intermediary metabolism). Likewise, bacterial resistance to antibiotics is multifaced. Resistance may involve decreased entry of the drug, changes in the receptor (target) of the drug, metabolic inactivation of the drug or production of "dummy" targets to disable the drug [7]. The resistance to the drug can be cooperative processes. The most compelling example is metabolic inactivation of the drug done outside the cells – a significant reduction in the drug's concentration can be done only with high concentration of the inactivating agent which requires dense population of bacteria. In many other cases, it is known experimentally that bacteria in an organized colony are much more resistant to antibiotics than the same bacteria in a suspension. This leads to our suggestion that impairing the cooperation of bacteria might render them as sensitive to antibiotics as they are in suspension.

Cooperation in the face of harsh condition is quite common among bacteria. The view of bacteria as unicellular microbes, a collection of non-interacting, identical passive "entities", persisted for a long time. Only during the last decade has a new approach emerged, one in which bacteria are sentient, interactive organisms, capable of sophisticated collective activity [8,9,10,11]. The cooperative behavior is observed in a diverse range of bacteria, which possess an unexpectedly broad repertoire of chemical and physical signaling mechanisms. *Escherichia coli*, gut bacteria, and *Salmonella typhimurium* use attractive chemotaxis call their peers for aid in detoxification of their environment [12,13,14]. *Bacillus subtilis*, *Paenibacillus dendritiformis* and *Serratia marcescens* cooperatively produce lubrication fluid in which they swim [15,16]. *P. dendritiformis* use repulsive chemotactic signaling to report their peers about unfavorable locations [17,18,19]. *Bacillus circulans* and *Paenibacillus vortex* utilize cell-cell physical interaction for response to external signals [20]. Thus the bacteria can communicate and organize their colonies into multicellular aggregates with new properties. Such aggregate can be viewed as a community or even as a multicellular organism [8,10].

The bacterial colonies, utilizing the bacteria's cooperative capabilities, develop complex spatio-temporal patterns in response to adverse growth conditions. In the past we have studied the response of bacterial colonies to three types of the external environmental stress: food shortage (starvation), hard media (limitations on movement) and indirectly, waste overload (where the bacteria have to decompose toxins they created themselves) [17,21,22,23,20]. For this subject we invoked insights and mathematical tools from previous studies of pattern formation in non-living systems and using generic modeling approach. Thus we were able to reveal novel bacterial strategies that account for the salient fea-

tures of the observed patterns. Using the models, we have demonstrated how bacterial chemo-communication (specifically, we focused on attractive and repulsive chemotactic signaling) leads to colonial self-organization that can only be achieved via cooperative behavior of the cells.

In this paper we present the response of the colonial growth dynamics to the presence of antibiotics in various stressful conditions (For further details see [24]). In medicine, resistance to antibiotics is usually considered a qualitative property: either a specific strain of bacteria is resistant to a specific antibiotic or it is not. The bacteria are generally regarded as being 'resistant' if they can tolerate the maximal concentration of the antibiotic which is non-toxic to the treated humans or animals. However, it is known that bacterial colonies are often more resistant than the individual cells, thereby blurring the borderline between resistance and sensitivity. Moreover, antibiotics affect cells differently depending on the growth and the physiological state of the cells. Our results show that susceptibility to antibiotic should be regarded as quantitative property: the bacteria react to antibiotic even in concentrations far below the critical effective concentrations that stop their growth.

Pathogenic bacteria form 2D films on many systems such as the cornea and sclera of the eye, in the urinary system, in the respiratory system (especially the lungs), on medical devices, etc. In this paper we study the response of the colonial growth dynamics to the presence of antibiotics under various stressful conditions, when the colony is grown on a surface. We have models which can describe the growth dynamics of different bacteria strains under various growth conditions [17,23,19,20,25]. From comparison between the simulations of the models and the experimental observations we can gain insights about the response of the bacteria to the antibiotic stress. From our existing model of chiral growth, for example, we conjecture that the changes in the colonial patterns caused by the antibiotic co-trimoxazole (Septrin, trademark of Wellcome) are due to elongation of the bacterial cells. This is consistent with the method of action of Septrin, which interferes with bacterial metabolism and might mimic starvation conditions – conditions under which the bacteria are usually longer.

2 Bacterial Strains

For the experiments described in this paper we use three groups of bacterial strains, all of which we initially isolated from plates containing *Bacillus subtilis* [26,27,28,29]. The strains were subsequently identified as belonging to the genera Paenibacillus. The strains display characteristic and distinct colonial morphologies and thus are characterized as three different morphotypes. We define morphotype as a group of bacteria exhibiting typical colonial patterns. Two strains belong to the same morphotype if under similar growth conditions they develop colonies of the same morphology for a range of growth conditions. Colonial growth patterns of a morphotype are inheritable and can be generated following inoculation with a single cell [11,30] (note that this definition does not exclude the possibility that bacteria from different species will belong to the same

morphotype, or that a specific species can belong to two different morphotypes depending on conditions).

Morphologies and studies of $\mathcal{T}$ morphotype are presented in Refs [26,31]. A representative strain from $\mathcal{T}$ morphotype was identified in [28] as *Paenibacillus dendritiformis* strain T168. Colonial patterns of the $\mathcal{T}$ morphotype on semi-solid agar are characterized by branched patterns with the branching growing in tip-splitting manner.

Morphologies and studies of $\mathcal{C}$ morphotype (for the *P. dendritiformis*) are presented in Refs [27,23,32]. When grown on semi-solid agar, colonies of bacteria from the $\mathcal{C}$ morphotype typically have thin branches, all twisted with the same handedness. New branches are side-branches of existing branches.

Morphologies and studies of $\mathcal{V}$ morphotype are presented in Refs [27,33,34,20]. A representative strain from $\mathcal{V}$ morphotype was identified in [29] as *P. vortex* strain V168. When grown on hard agar, bacteria of this morphotype organize themselves into vortices. Colonial patterns of these bacteria are branchy, where each branch is led by a droplet of cells. Each leading droplet consists of up to million cells that rotate around a common center (hence the term vortex).

3 The Effect of Antibiotics on Colonies of the $\mathcal{T}$ Morphotype

In most of our experiments we have studied the effects of the antibiotic co-trimoxazole (Septrin), usually at a level of 0.0003% ($3\mu g/ml$). Three phenomena are observed: 1. The rate of colonial expansion is lower. 2. The growth becomes much more ramified, i.e. there is a decrease in the fractal dimension. 3. For growth on harder surfaces in the presence of Septrin, there is more pronounced weak chirality.

In Fig. 1 we demonstrate the second phenomenon. For these growth conditions, the fractal dimension D_f changes from $D_f = 2$ to $D_f \approx 1.8$ in the presence of Septrin, while the growth velocity changes only slightly. Both weak chirality and a decrease in fractal dimension are observed in Fig. 2. Note that in the presence of Septrin the velocity (colonial expansion rate) is reduced to less than half of that of the growth without antibiotics.

The antibiotic effect is weaker as the peptone level is increased. At high peptone levels and hard substrates (2% agar and above) we frequently observe (in more than half of the colonies) bursts of sectors with a faster growth (Fig. 3). The new sectors are apparently produced by mutants which grow better in the presence of Septrin. Picking experiments demonstrate that the sectors are indeed connected with inheritable genetic features. The frequency of bursts of sectors is also increased when the level of Septrin is increased.

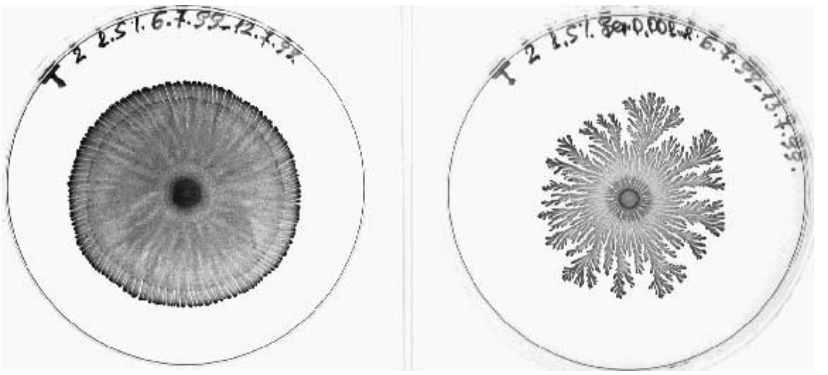

Fig. 1. The effect of Septrin on the fractal dimension of the $\mathcal{T}$ morphotype (left: without antibiotic, right: with $3\mu g/ml$ Septrin). $\mathcal{T}$ morphotype grown at $2g/l$ peptone and 1.5% agar. Growth times are 6 days (left) and 7 days (right).

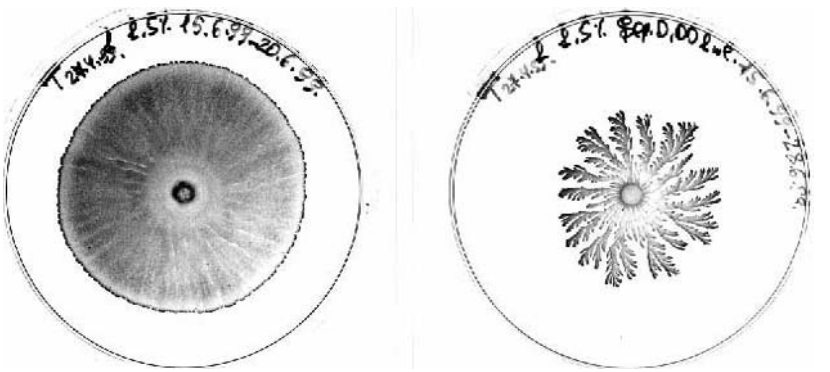

Fig. 2. Similar to Fig. 2.1, for growth at $1g/l$ peptone. Growth times are 5 days (left) and 10 days (right).

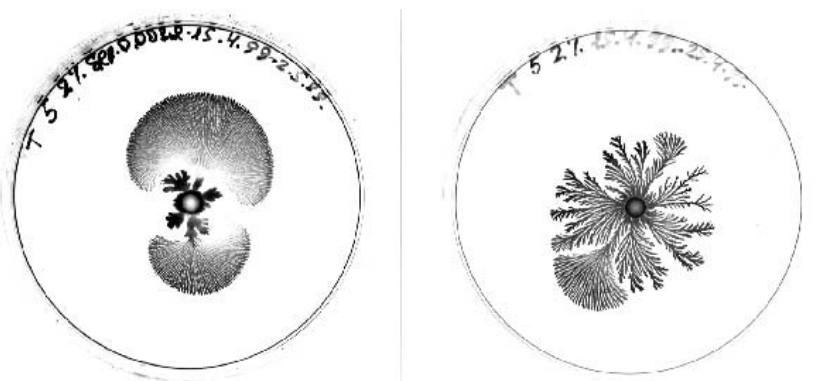

Fig. 3. Examples of bursts of sectors during growth at $5g/l$ peptone and 2% agar. Growth times are 17 days (left, $3\mu g/ml$ Septrin) and 10 days (right, without antibiotics).

4 The Effect of Antibiotics on Colonies of the $\mathcal{C}$ Morphotype

Bacteria of the $\mathcal{C}$ (chiral) morphotype belong to the same species as bacteria of the $\mathcal{T}$ morphotype, yet their colonial growth patterns are very different [23,28,32]. As with $\mathcal{T}$ morphotype, we study mainly the effect of the antibiotic Septrin on the growth dynamics.

The main feature of the growth pattern of the $\mathcal{C}$ morphotype is the thin branches which are all twisted with the same handedness (strong chirality). The effect of Septrin on the growth is demonstrated in Figs. 4 and 5. The effect of other types of antibiotics on two strains of the $\mathcal{C}$ morphotype is shown in Figs. 6 and 7.

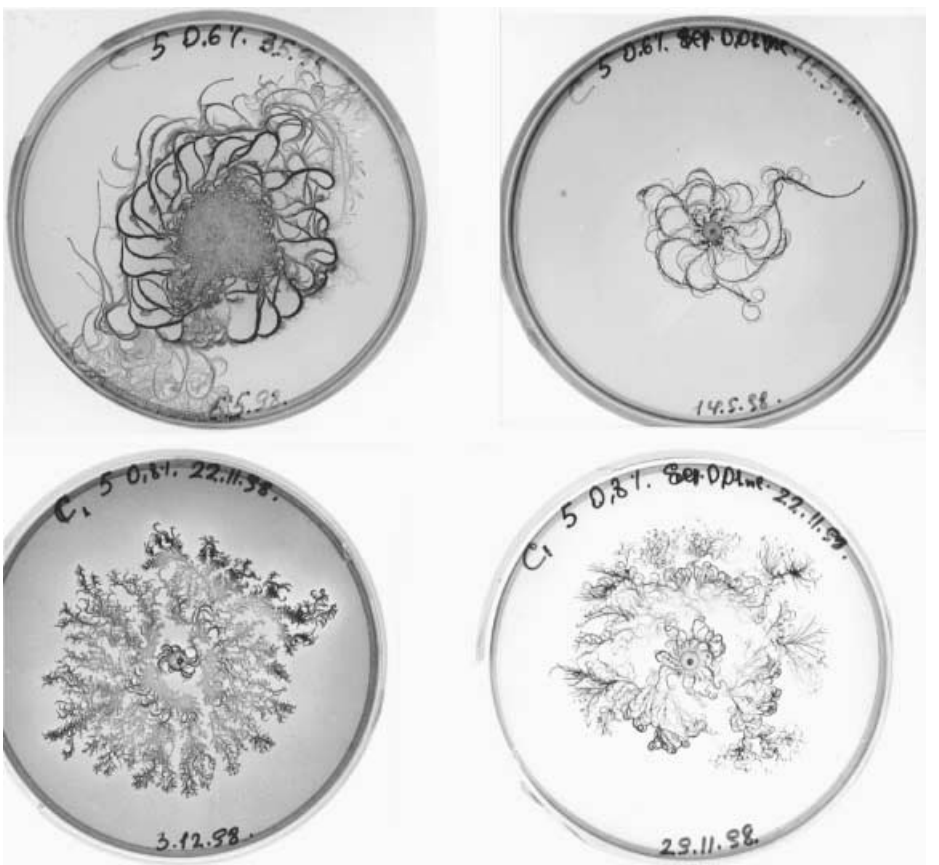

Fig. 4. Typical colonies of $\mathcal{C}$ morphotype grown with (right) and without (left) Septrin. Septrin concentration $30\mu g/ml$. The colonies were grown on media of $5g/l$ peptone, 0.6% agar concentration (top) and 0.8% agar concentration (bottom). Growth times are 3 days (top left and right),11 days (bottom left), 7 days (bottom right).

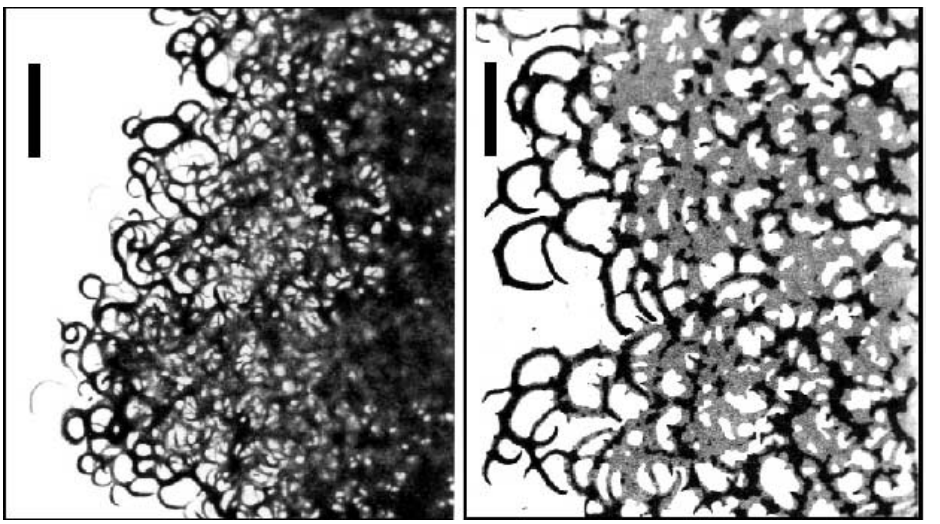

Fig. 5. Closer look at colonies of $\mathcal{C}$ morphotype reveals the twisted branches and the effect Septrin has on them. The colonies are grown on media of $2g/l$ peptone and 1.25% agar concentration. On the left: a colony grown without Septrin. The average radius of curvature is $\sim 1.0mm$ (the branches at the edge of the colony are the most indicative, as the pattern is not obscured by over-growth with other branches). Bar size is $5mm$. On the right: a colony grown with Septrin at a concentration of $30\mu g/ml$. The average radius of curvature is $\sim 1.8mm$. Bar size is $5mm$.

5 The Effect of Antibiotic on Colonies of $\mathcal{V}$ Morphotype

There is also a pronounced effect of Septrin on the growth patterns of the $\mathcal{V}$ morphotype. In Fig. 8 we show a typical example. Septrin causes the pattern to become much more radially oriented and eliminates most of the side branches.

6 Interpretation of the $\mathcal{T}$ Morphotype Observations within the Framework of the Non-linear Diffusion Model

6.1 The Non-linear Diffusion Model

In the context of branching growth of bacterial colonies, the continuous modeling approach has been pursued recently by Mimura and Matsushita *et al.* [35,36], Kawasaki *et al.* [37], Kitsunezaki [38], Ben-Jacob *et al.* [32,39,25,40] and Lacasta *et al.* [41]. In Ref. [25] we present a summary and critique of this approach.

We have previously shown [39] that the Non-Linear Diffusion (NLD) model [38,42] provides a good approximation to the more detailed Lubricating bacteria model. Hence this is the model we use.

Fig. 6. The effect of antibiotics on the growth of $\mathcal{C}$ morphotype strain CB2. Growth conditions are $5g/l$ peptone and 0.6% agar concentration. Top left: in the absence of antibiotics (growth time 6 days). Top right: The effect of Chloramphenicol. Middle row: The effect of 3 $\mu g/ml$ Septrine (growth time 6 days). Left and right are the same growth conditions. There are differences in the patterns, resulting from variations in the number of bacteria in the inoculating droplet. Bottom row: 30 $\mu g/ml$ Septrin (growth time 6 days). Again there are differences between the two patterns. Note that although the level of Septrin is 10 times higher than in the middle row, the growth is faster. We assume this results from the fact that the antibiotic causes the bacteria to become longer, and there is also a stronger response to repulsive chemotaxis.

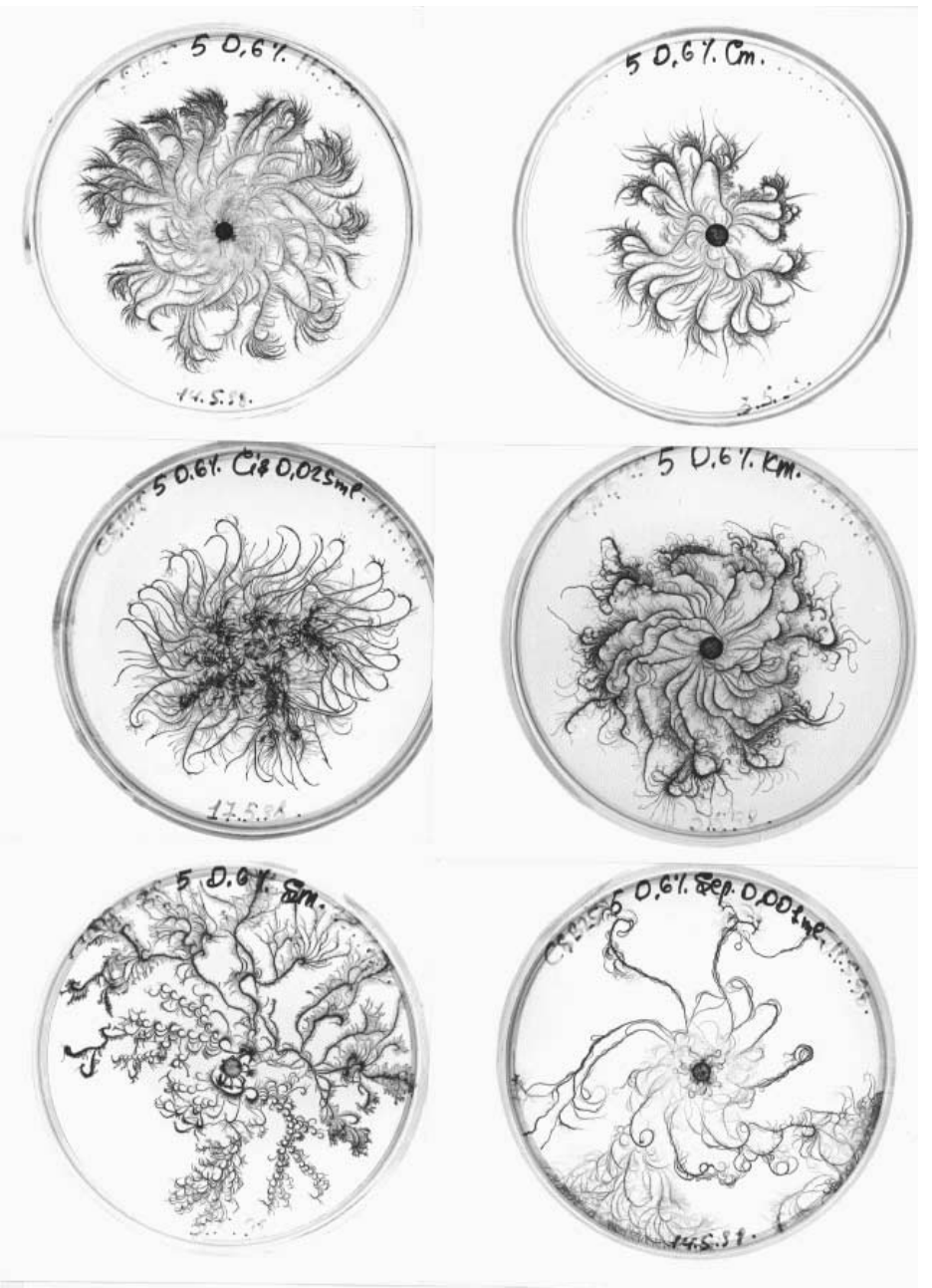

Fig. 7. The effect of different kinds of antibiotics on the growth of $\mathcal{C}$ morphotype strain CSB25. Growth conditions are $5g/l$ peptone and 0.6% agar concentration. Top left: No antibiotics (growth time 3 days). Top right: Chloramphenicol (6 days). Middle left: Abiplatin, a chemo-therapeutic material (3 days). Middle right: Kanamycin (6 days). Bottom left: Streptomycin (6 days). Bottom right: Septrin (3 days).

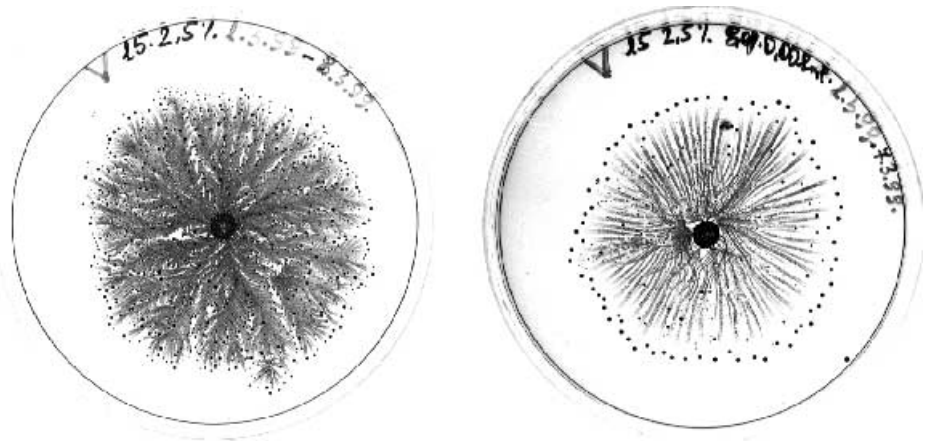

Fig. 8. The effect of Septrin on the growth pattern of the $\mathcal{V}$ morphotype for Growth conditions of $15g/l$ peptone and 2.5% agar. Growth times are 7 days (left, no antibiotics) and 6 days (right, $3\mu g/ml$ Septrin). Note that the growth is faster under the pressure of the antibiotic.

The time evolution of the 2D bacterial density $b(\mathbf{x}, t)$ is given by [25]:

$$\frac{\partial b}{\partial t} = \nabla(D(b)\nabla b) + \beta n b - \mu b \tag{1}$$

The first term on the RHS describes the bacterial movement, with $D(b) = D_0 b^k$ (where D_0 and $k > 0$ are constants). The second term reflects the population growth, which is proportional to food consumption. $n(\mathbf{x}, t)$ is the 2D nutrient concentration and β the nutrient-to-bacteria conversion factor. The third term describes the process of bacterial transformation into stationary, pre-spore state, with μ being the relevant rate. The time development of the nutrient concentration $n(\mathbf{x}, t)$ and the density $s(\mathbf{x}, t)$ of stationary bacteria are given by:

$$\frac{\partial n}{\partial t} = D_n \nabla^2 n - \beta b n \tag{2}$$

$$\frac{\partial s}{\partial t} = \mu b \tag{3}$$

A typical pattern generated by this model is depicted in Fig. 9.

6.2 Metabolic Load

Colonies of $\mathcal{T}$ morphotype can grow in the presence of Septrin, i.e. they are resistant to the concentrations of the antibiotic normally used in the experiments (Sec. 3). We do not know the molecular basis of this resistance. It seems plausible that at least one of the effects is an increase in the metabolic load of the bacteria. We would like to check whether such effect exists, and if so, are there any other effects (i.e. is it sufficient to explain the observations?) In the NLD model

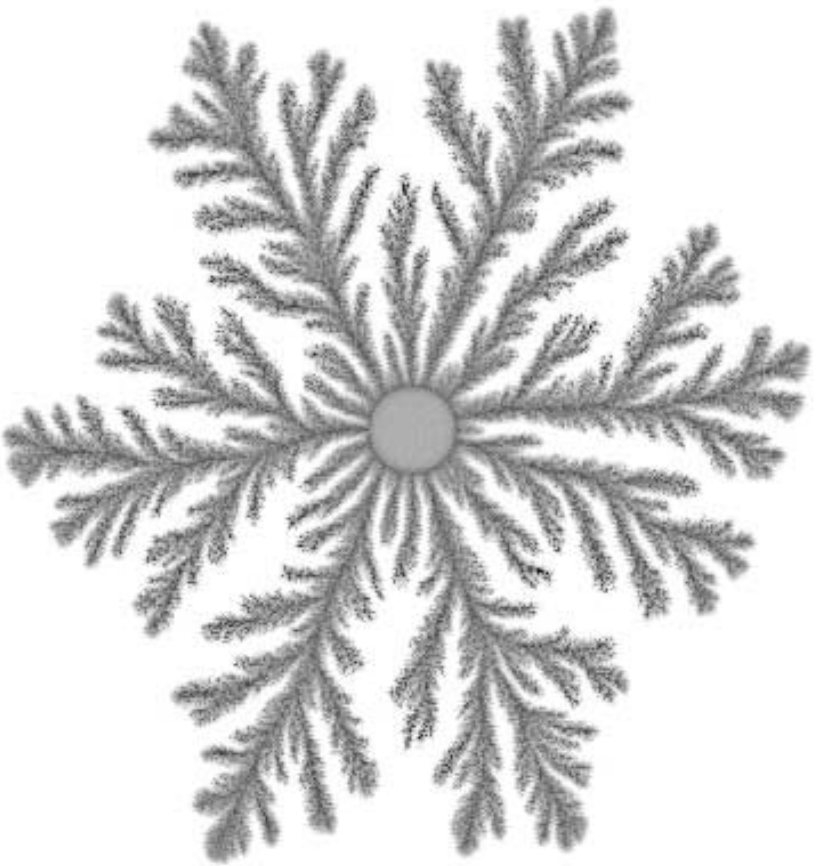

Fig. 9. 2D growth pattern ($b+s$) of the NLD model. The parameters are: $D_0 = 0.1, k = 1, \beta = 1.0, \mu = 0.05, n_0 = 1$. Time of growth: 1.676 (dimensionless units)

increased metabolic load is captured by an increase in μ. As we have remarked previously [39], μ can also be interpreted as the relative amount of energy the bacteria require for their life-sustaining activities, energy which is subtracted from the amount available for their reproduction. Our assumption is that the antibiotic stress forces the bacteria to invest more energy in maintenance and hence corresponds to an increased μ.

As shown in Figs. 10 and 11 increasing μ leads to thinner, ramified branches and to a lower growth velocity, in agreement with the experimental observations.

6.3 The Effect of Chemotaxis

Chemotaxis is a bias of movement according to the gradient of a chemical material. Chemotactic signaling is a chemotactic response to a chemical emitted by

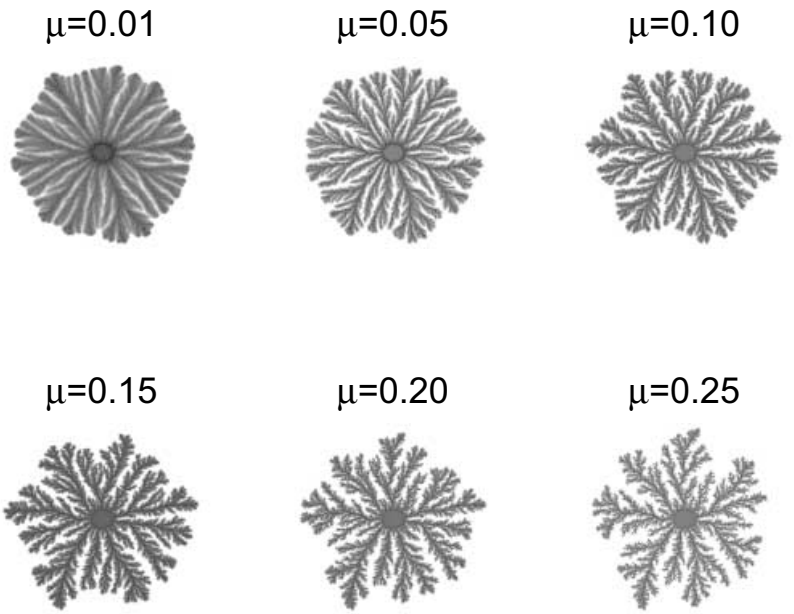

Fig. 10. Growth patterns of the NLD model, with varying values of μ. All other parameters are as in Fig. 9. Time of growth is chosen for each set of growth conditions so that all patterns will have approximately the same size.

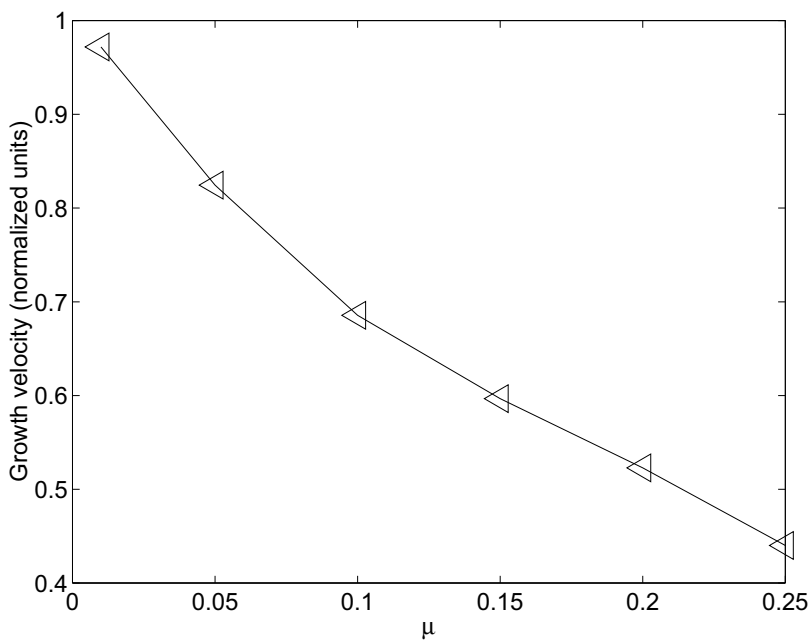

Fig. 11. Growth velocity as a function of μ. All other parameters are as in Fig. 9.

the bacteria. We have previously postulated [14] that bacteria employ repulsive chemotactic signaling under conditions of very low nutrient level, i.e. under nutritional stress. It was of interest to determine whether under antibiotic stress, repulsive chemotactic signaling will also play a role.

We have incorporated the effects of chemotaxis in the continuous model by introducing a *chemotactic flux* $\boldsymbol{J}_{chem}$, which is written (for the case of a chemorepellent and a linear diffusion) as [43]:

$$\boldsymbol{J}_{chem} = -b\chi(r)\nabla r \tag{4}$$

where $r(\mathbf{x}, t)$ is the concentration of a chemorepellent and $\chi(r)$ is the chemotactic sensitivity to the repellent, given usually by the "receptor law" [43]. In addition, we write an equation describing the diffusion, the production and decomposition of the chemorepellent.

As seen in Fig. 12, the main effects of repulsive chemotaxis are to make the branches radially oriented, and to increase the colony expansion. We expect that the combined effect of both repulsive chemotaxis and an increase in μ is to produce slower and more radially oriented colonial branches, again in keeping with the observations.

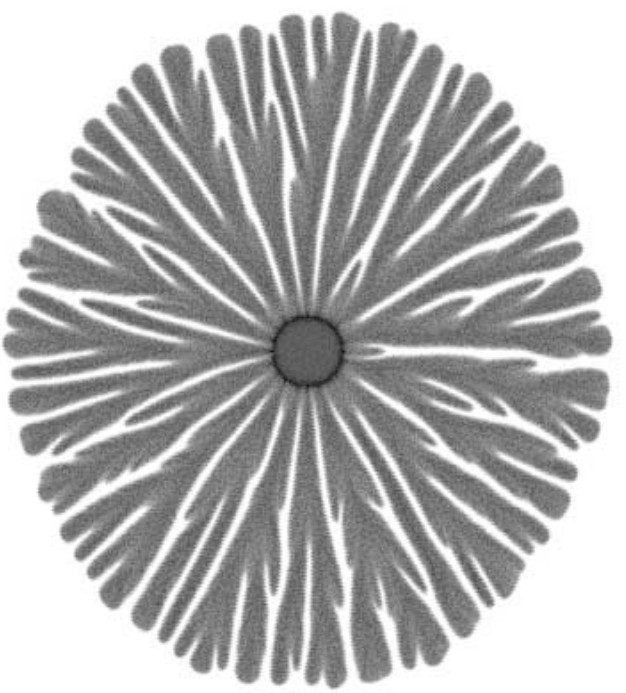

Fig. 12. 2D growth pattern $(b + s)$ of the model with repulsive chemotactic signaling included. $\chi_0 = 1.0$. Other parameters as in Fig. 9.

6.4 The Appearance of Weak Chirality

Colonies of $\mathcal{T}$ morphotype grown on hard agar surfaces exhibit branching patterns exhibiting a global twist with the same handedness. We refer to such growth patterns as having weak chirality.

Ben-Jacob *et al.* [23] proposed that the high viscosity of the "lubrication" fluid during growth on a hard surface limits the free rotation of bacteria while tumbling. It was further assumed that the rotation should be relative to a specific direction. The gradient of a repulsive chemotactic signaling field (which is long-ranged and radial) was assumed to define the required specified direction.

In Ref. [32] we obtain weak chirality in the reaction-diffusion model by modifying the chemotactic mechanism and causing it to twist: We alter the expression for the chemotactic flux $\boldsymbol{J}_{chem}$ so that it is no longer oriented parallel to the chemical gradient (∇R). Rather, $\boldsymbol{J}_{chem}$ is oriented in the direction of a rotated vector $\hat{\mathbf{R}}(\theta)\nabla R$, where $\hat{\mathbf{R}}(\theta)$ is the two-dimensional rotation operator and θ is the rotation angle. The effect of rotating the repulsive chemotaxis, as depicted in Fig. 13, is to introduce weak chirality, with the degree of chirality determined by the rotation angle θ.

The phenomenon of weak chirality seems to be enhanced by the presence of antibiotics. This may stem from the elongation of individual cells under the influence of antibiotics, elongation which increases the orientational interaction between cells and limits their rotation during tumbling.

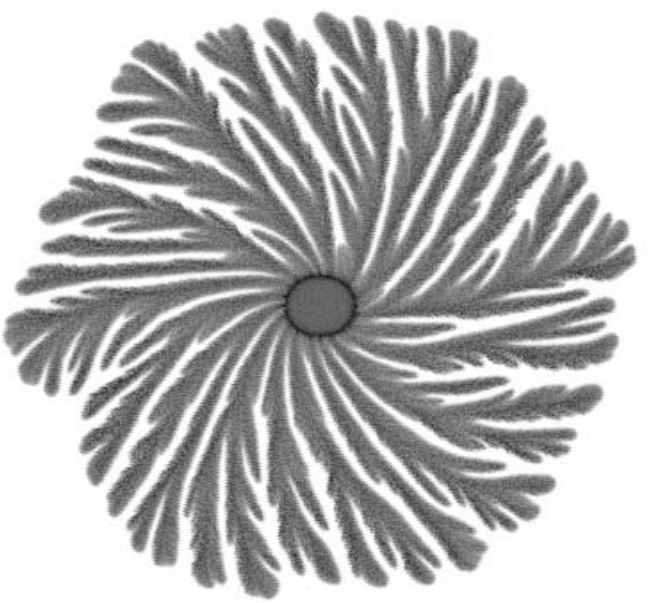

Fig. 13. Growth patterns of the NLD model with a "squinting" repulsive chemotactic signaling, leading to weak chirality. Parameters as in Fig. 12.

6.5 Sector Formation in Expanding Colonies

The phenomenon of bursts of new sectors of mutants during the growth of bacterial colonies is well known [44,45]. If the mutants have the same growth dynamics as the parental, wild-type organism, they will usually go unnoticed (unless some property such as pigmentation distinguishes them). If, however, the mutants exhibit a different growth dynamics, the appearance of a distinct sector with a different growth pattern might indicate their presence.

In Ref. [46], we generalized the NLD model to study bursts of sectors. This is done by introducing two fields, one for the densities of the wild-type bacteria ("type 1") and the other for the mutants ("type 2"), and allowing some probability of transition from wild-type to mutants.

Following our experimental observations and the numerical studies, we are able to appreciate what factors – geometrical, regulatory and others – favor the segregation of the mutant population. These factors include the expansion of the colony, the branching of the pattern, the effect of chemotaxis, and specific advantages possessed by the individual mutants. Fig. 14 shows an example of our simulations, for the case in which the mutant exhibits a stronger response to repulsive chemotactic signaling.

A high level of antibiotics, like other forms of stress, seems to encourage the appearance of new mutations [47], so that many sectors emerge in the developing colony.

6.6 The Secretion of an Inhibitory Substance During Colonial Embryonic Stage

For a more critical test of the effect of Septrin, we have studied growth patterns in which the initial inoculation is not at a point in the center of the Petri dish, but rather along a line, along two parallel lines, a circle, or a line together with a circle.

First, we show in Fig. 15 the effect on growth started from a single line. The main effect in this case (in the presence of Septrin) is the enhanced growth from the two poles and the reduced growth from the center. In Fig. 16 we show growth from inoculation along two parallel lines. In the absence of antibiotics there is internal growth (between the lines) when the lines are far apart, and there is no internal growth when the lines are closer.

We suggest that that the absence of growth does not result from depletion of food. We propose that it might result from accumulation of inhibitory substances secreted by the bacteria during the initial stage between inoculation and the first burst from the inoculum – a period of several hours. In the presence of antibiotics there is no internal growth for both cases. This might indicate that in the presence of antibiotics there is a higher emission of inhibitory substances by the bacteria.

Indeed, simulations of the NLD model lead to growth along the inoculation line, as well as between the two inoculation lines. The inclusion of repulsive chemotaxis does not change this fact (Fig. 17). Thus, we have included in the

Fig. 14. Mutant with a higher sensitivity to repulsive chemotactic signaling, in a branching colony: Numerical simulation of the NLD model. The mutant erupts in a fan-like sector from the colony of wild-type bacteria. The model parameters for the wild-type: $D_0 = 0.1, k = 1, \beta = 1.0, \mu = 0.2, \chi_0 = 0.5$, mutant has $\chi_0 = 1.0$.

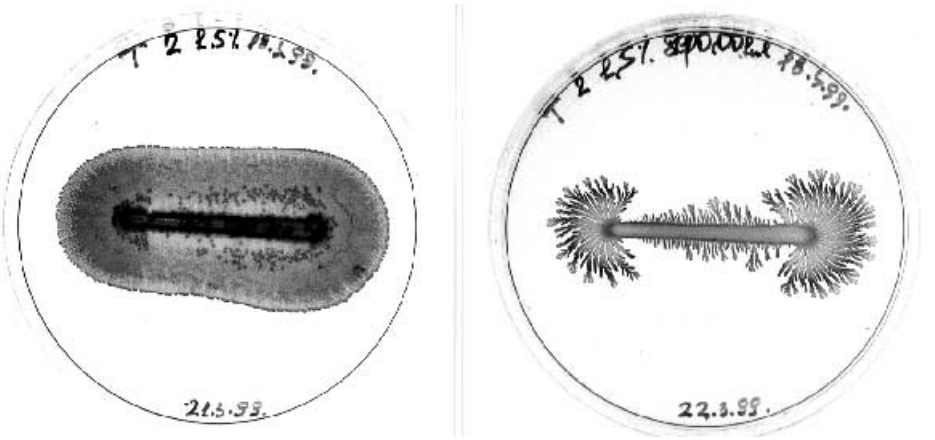

Fig. 15. Growth from a line inoculum, for $2g/l$ peptone and 1.5% agar. Growth times are 3 days (left, no antibiotics) and 4 days (right, $3\mu g/ml$ Septrin).

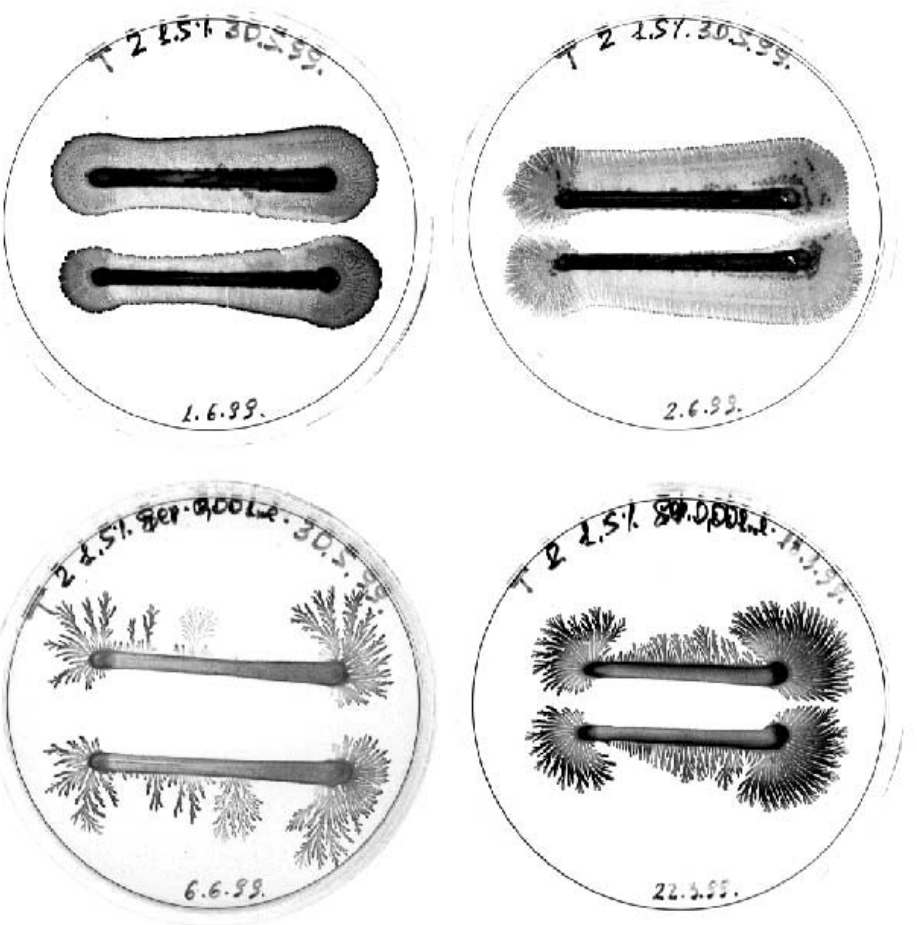

Fig. 16. Growth from two line inocula, for $2g/l$ peptone and 1.5% agar. Top: no antibiotics. Bottom: $3\mu g/ml$ Septrin. Left: larger distance between the lines ($17mm$ gap on the left, $10mm$ gap on the right). Growth times: Top left 3 days, Top right 4 days, Bottom left 8 days, Bottom right 7 days.

model the secretion of an inhibitory diffusive chemical during the initial stage of colonial development. In the case of a symmetric circular development from a droplet inoculation at the center of the Petri dish, the emission of the inhibitory material only leads to a delay in the expansion of the colony. There is also some small change in the colonial pattern during the initial growth, but it is hard to quantify. The effect is much more dramatic for inoculation along a line or two parallel lines, as shown in Figs. 18 and 19. There is a strong effect of reduction of the internal growth. When the level of inhibitory material is high enough, the internal growth is completely eliminated.

After bacteria are inoculated on semi-solid agar or a solid surface, it takes several hours until the colony starts to expand outwards. We refer to this period as the colonial embryonic stage, or the consolidation period. Although many interesting phenomena might take place during this stage, it has rarely been studied in detail. In addition to the secretion of an inhibitory substance during this time, the isolated bacteria also develop their "chemical identity" as a specific colony. That is, bacteria which belong to a specific colony are capable of identifying their peers. Two separate colonies which grow toward each other will not fuse, even if inoculated from the same source. However, if we place a barrier in front of an expanding colony, so the barrier in engulfed from both sides, the two fronts will fuse. We believe that the emergence of the colonial identity is a most important process that should be further studied.

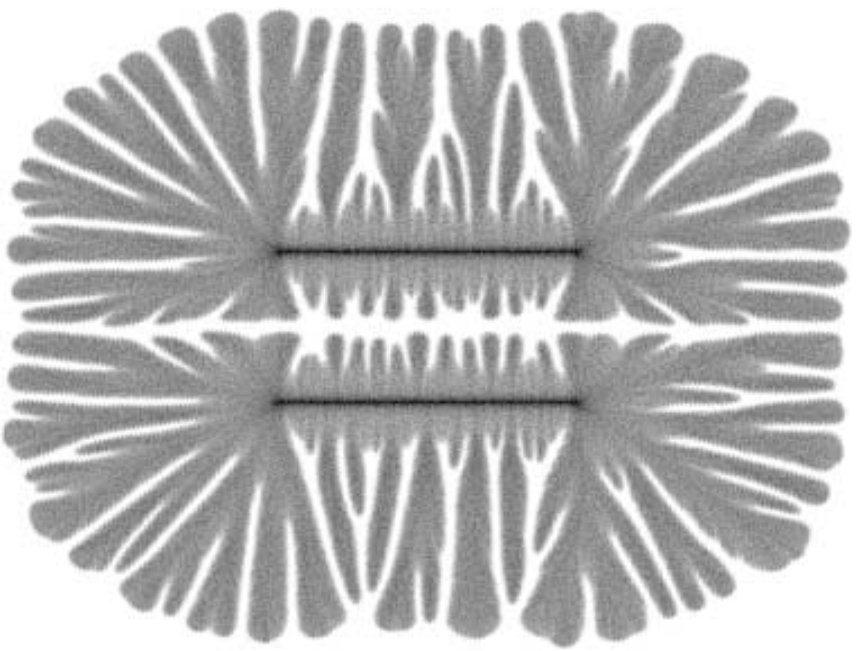

Fig. 17. A colony from two lines of inoculation, in the presence of a repulsive chemotactic field. The parameters are as in Fig. 9.

7 Modeling the Chiral Growth of $\mathcal{C}$ Morphotype

The Communicating Spinors (CS) model [32] was developed to explain the chirality of the $\mathcal{C}$ morphotype colonies. We showed that handedness of flagella structure and dynamics of rotation are amplified to affect the macroscopic pattern, leading to the observed chirality. It does so in the same manner in which crystalline anisotropy leads to the observed symmetry of snowflakes [48].

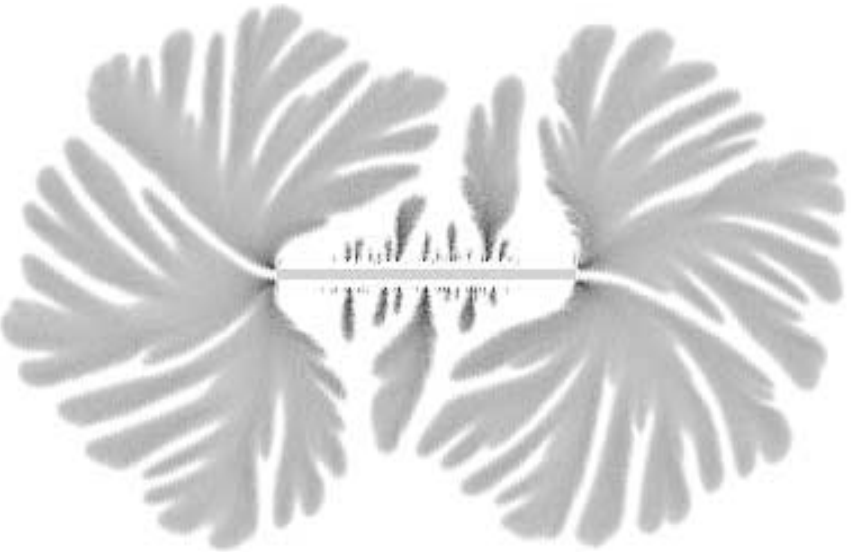

Fig. 18. A colony developing from a line inoculation, in the presence of a repulsive chemotactic field (leading to weak chirality), and a field of inhibitory substance. NLD model parameters are as in Fig. 9.

It is known [49,50,51] that flagella have specific handedness of structure and rotation dynamics. Ben-Jacob *et al.* [23] proposed that the latter is the origin of the observed chirality. In a fluid, as the flagella unfold, the cell tumbles and ends up at a new random angle relative to the original one. The situation changes for quasi 2D motion – motion in a "lubrication" layer thinner then the cellular length. We assume that in this case, of swimming in a plane, the tumbling has a well defined handedness of rotation. Such handedness requires, along with the chirality of the flagella, that the cells can distinguish "up from down". The growth in an inverted petri-dish shows the same chirality as growth in a plate "right-side up". Therefore, we think that the determination of up *vs.* down is done either via the vertical gradient of the nutrient concentration, via the vertical gradient of signaling materials inside the substrate, or via the adhesion of the cells to the surface of the agar. The latter is the most plausible alternative, as on soft enough agar which enables the bacteria to swim *below* the surface of the agar, the handedness of the branches is reversed [32].

In Fig. 20 we show results of simulation of the CS model (See [24] for more details).

8 Conclusions

In this paper we showed the effect of non-lethal levels of antibiotic material on bacterial cooperative organization. We presented experimental evidence of the changes in colonial development of *Paenibacillus dendritiformis* and *P. vortex*

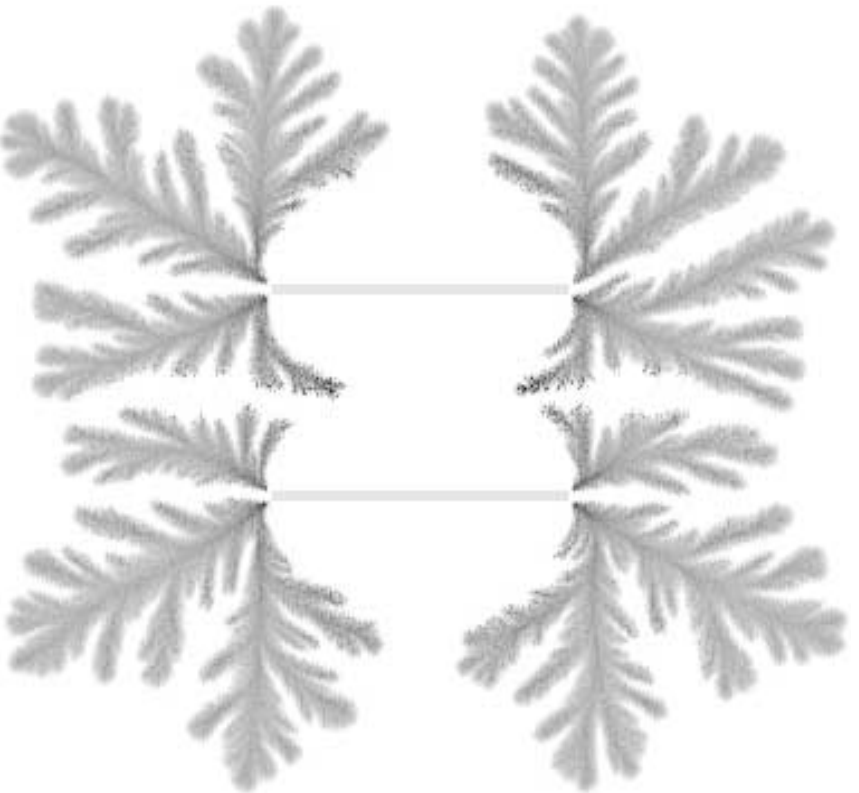

Fig. 19. A colony from two lines of inoculation, in the presence of an inhibitory substance. The NLD model parameters are as in Fig. 9.

caused by sub-lethal concentrations of antibiotics. We focused on the effect of Septrin (co-trimoxazole) on the colonial organization of *P. dendritiformis* .

Various changes in the colonial growth patterns appeared in response to exposure to Septrin. Usually (but not always) the colonial growth velocity was reduced. Branching, tip-splitting patterns of *P. dendritiformis* were affected by reduction in the colonial fractal dimension, appearance of pronounced weak chirality and pronounced radial orientation of the growth. The radius of curvature of twisted branches in chiral patterns was increased. The growth patterns of colonies of *P. vortex* were affected in a very complex manner. One of the most noticeable feature is that the colonial growth velocity was sometimes *increased* in response to antibiotic.

Numerical studies of various generic models indicate the causes for modifications in the patterns. From these studies we conclude that the *P. dendritiformis* cells responded to Septrin in a multilayered fashion. The bacteria as individuals increased their length. The rate of bacterial reproduction was reduced due to metabolic load. The colonial organization was altered apparently due to adjustment in the chemotactic signaling between the bacteria.

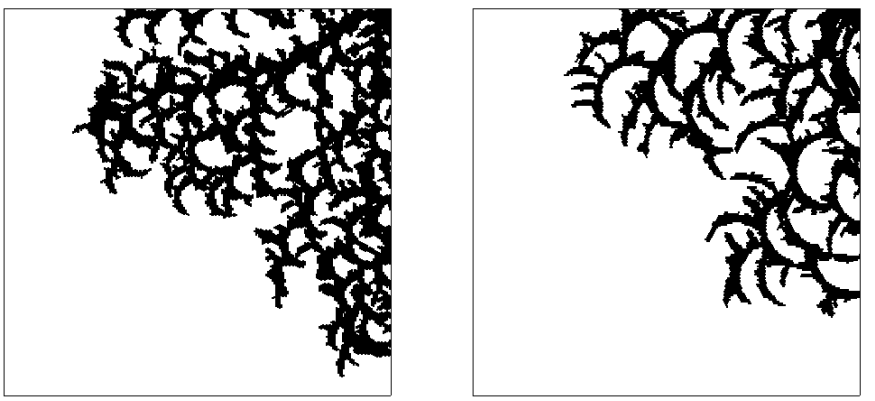

Fig. 20. Colonies of $\mathcal{C}$ morphotype modeled by the CS model. A decrease in the values of η and ϕ_c leads to an increase in the average curvature of the branches. On the left $\eta = \phi_c = 7^\circ$. On the right $\eta = \phi_c = 3^\circ$. The decrease in the parameters' value represents elongation of the bacterial cells.

The theoretical studies presented here are just preliminary results and only demonstrate the complexity introduced into the biological system by exposure to antibiotic. Much effort is needed to reveal the affects of other antibiotics on colonial organization and to understand the differences and similarities between the effects of different materials. The effects of antibiotic on colonies of *P. vortex* are yet to be understood. We hope that the results will enable us to generalize the conclusions to other bacteria or other antibiotic materials.

Acknowledgments

The results presented here are based on a joint project with Inon Cohen, David L. Gutnick, Marianna Tcherpakov, Dirk Helbing and Ilan G. Ron.

We are grateful to Inna Brainis and Rina Avigad for technical assistance and to Amikam Shoob for photography of the colonies. This research has been partially supported by a grant from the Israeli Academy of Sciences, grant no. 461/98, by the Israeli-US Binational Science Foundation BSF, grant no. 00410-95 and by a grant from IMK Almene Foundation. One of us (EBJ) wishes to thank the organizers of the XVII Sitges Conference for their hospitality during the meeting.

References

1. C. F. Amáblie-Cueva, M. Cárdenas-García, and M. Ludgar. Antibiotic resistance. *Am. Sci.*, 83:320–329, 1995.
2. P. Radetsky. Last days of the wonder drugs. *Discover*, 19, 1998.

3. S. B. Levy. The challenge of antibiotic resistance. *Sci. Am.*, March, 1998.
4. R. V. Miller. Bacterial gene swapping in nature. *Sci. Am.*, 278(1), 1998.
5. J. A. Shapiro. Natural genetic engineering in evolution. *Genetica*, 86:99–111, 1992.
6. R. Losick and D. Kaiser. Why and how bacteria communicate. *Sci. Am.*, February, 1997.
7. S. Baron, editor. *Medical Microbiology*. The University of Texas Medical Branch at Galveston, fourth edition, 1996.
8. J. A. Shapiro. Bacteria as multicellular organisms. *Sci. Am.*, 258(6):62–69, 1988.
9. D. Kaiser and R. Losick. How and why bacteria talk to each other. *Cell*, 73:873–887, 1993.
10. L.J. Shimkets and M. Dworkin. *Myxobacterial* multicellularity. In J. A. Shapiro and M. Dworkin, editors, *Bacteria as Multicellular Organisms*. Oxford University Press, New-York, 1997.
11. E. Ben-Jacob. From snowflake formation to the growth of bacterial colonies. part II: Cooperative formation of complex colonial patterns. *Contemp. Phys.*, 38:205–241, 1997.
12. E. O. Budrene and H. C. Berg. Complex patterns formed by motile cells of *escherichia coli*. *Nature*, 349:630–633, 1991.
13. Y. Blat and M. Eisenbach. Tar-dependent and -independent pattern formation by *Salmonella typhimurium*. *J. Bac.*, 177(7):1683–1691, 1995.
14. E. O. Budrene and H. C. Berg. Dynamics of formation of symmetrical patterns by chemotactic bacteria. *Nature*, 376:49–53, 1995.
15. M. A. Marahiel, M. M. Nakano, and P. Zuber. Regulation of peptide antibiotic production in *bacillus*. *Mol. Microbiol.*, 7:631–636, 1993.
16. J. D. Desai and I. M. Banat. Microbial production of surfactants and their commercial potential. *Microbiol. Mol. Biol. Rev.*, 61:47–64, 1997.
17. E. Ben-Jacob, O. Shochet, A. Tenenbaum, I. Cohen, A. Czirók, and T. Vicsek. Generic modeling of cooperative growth patterns in bacterial colonies. *Nature*, 368:46–49, 1994.
18. E. Ben-Jacob, O. Shochet, I. Cohen, A. Tenenbaum, A. Czirók, and T. Vicsek. Cooperative strategies in formation of complex bacterial patterns. *Fractals*, 3:849–868, 1995.
19. I. Cohen, A. Czirók, and E. Ben-Jacob. Chemotactic-based adaptive self organization during colonial development. *Physica A*, 233:678–698, 1996.
20. E. Ben-Jacob, I. Cohen, A. Czirók, T. Vicsek, and D. L. Gutnick. Chemomodulation of cellular movement and collective formation of vortices by swarming bacteria and colonial development. *Physica A*, 238:181–197, 1997.
21. E. Ben-Jacob, I. Cohen, O. Shochet, I. Aronson, H. Levine, and L. Tsimering. Complex bacterial patterns. *Nature*, 373:566–567, 1995.
22. L. Tsimring, H. Levine, I. Aranson, E. Ben-Jacob, I. Cohen, O. Shochet, and W. N. Reynolds. Aggregation patterns in stressed bacteria. *Phys. Rev. Lett.*, 75:1859–1862, 1995.
23. E. Ben-Jacob, I. Cohen, O. Shochet, A. Czirók, and T. Vicsek. Cooperative formation of chiral patterns during growth of bacterial colonies. *Phys. Rev. Lett.*, 75(15):2899–2902, 1995.
24. E. Ben-Jacob, I. Cohen, I. Golding, D.L. Gutnick, M. Tcherpakov, D. Helbing, and I.G. Ron. Bacterial cooperative organization under antibiotic stress. *Physica A*, 282(1-2):247–282, 2000.
25. I. Golding, Y. Kozlovsky, I. Cohen, and E. Ben-Jacob. Studies of bacterial branching growth using reaction-diffusion models of colonial development. *Physica A*, 260(3-4):510–554, 1998.

26. E. Ben-Jacob, H. Shmueli, O. Shochet, and A. Tenenbaum. Adaptive self-organization during growth of bacterial colonies. *Physica A*, 187:378–424, 1992.
27. E. Ben-Jacob, A. Tenenbaum, O. Shochet, and O. Avidan. Holotransformations of bacterial colonies and genome cybernetics. *Physica A*, 202:1–47, 1994.
28. M. Tcherpakov, E. Ben-Jacob, and D. L. Gutnick. *Paenibacillus dendritiformis* sp. nov., proposal for a new pattern-forming species and its localization within a phylogenetic cluster. *Int. J. Syst. Bacteriol.*, 49:239–246, 1999.
29. M. Tcherpakov, E. Ben-Jacob, I. Cohen, and D. L. Gutnick. *Paenibacillus vortex* sp. nov., proposal for a new pattern-forming species and its localization within a phylogenetic cluster. submitted to *Int. J. Syst. Bacteriol.*
30. E. Ben-Jacob, I. Cohen, and D. Gutnick. Cooperative organization of bacterial colonies: From genotype to morphotype. *Annu. Rev. Microbiol.*, 52:779–806, 1998.
31. E. Ben-Jacob, O. Shochet, A. Tenenbaum, I. Cohen, A. Czirók, and T. Vicsek. Communication, regulation and control during complex patterning of bacterial colonies. *Fractals*, 2(1):15–44, 1994.
32. E. Ben-Jacob, I. Cohen, I. Golding, and Y. Kozlovsky. Modeling branching and chiral colonial patterning of lubricating bacteria. In *Proceedings of 1998 IMA workshop: Pattern Formation and Morphogenesis*. Springer-Verlag, Berlin, 1999. (in press).
33. E. Ben-Jacob, O. Shochet, A. Tenenbaum, and O. Avidan. Evolution of complexity during growth of bacterial colonies. In P. E. Cladis and P. Palffy-Muhoray, editors, *Spatio-Temporal Patterns in Nonequilibrium Complex Systems*, Santa-Fe Institute studies in the sciences of complexity, pages 619–634. Addison-Weseley Publishing Company, 1995.
34. A. Dukler. Isolation and characterization of bacteria growing in patterns on the surface of solid agar. Senior project under the supervision of D. Gutnick., Tel-Aviv University, 1993.
35. M. Mimura, H. Sakaguchi, and M. Matsushita. A reaction-diffusion approach to bacterial colony formation. *preprint*, 1997.
36. M. Matsushita, J. Wakita, H. Itoh, I. Rafols, T. Matsuyama, H. Sakaguchi, and M. Mimura. Interface growth and pattern formation in bacterial colonies. *Physica A*, 249:517–524, 1998.
37. K. Kawasaki, A. Mochizuki, M. Matsushita, T. Umeda, and N. Shigesada. Modeling spatio-temporal patterns created by bacillus-subtilis. *J. Theor. Biol.*, 188:177–185, 1997.
38. S. Kitsunezaki. Interface dynamics for bacterial colony formation. *J. Phys. Soc. Jpn*, 66(5):1544–1550, 1997.
39. Y. Kozlovsky, I. Cohen, I. Golding, and E. Ben-Jacob. Lubricating bacteria model for branching growth of bacterial colonies. *Phys. Rev. E*, 59(6):7025–7035, 1999.
40. I. Cohen, I. Golding, Y. Kozlovsky, E. Ben-Jacob, and I. G. Ron. Continuous and discrete models of cooperation in complex bacterial colonies. *Fractals*, 7:235–247, 1999.
41. A. M. Lacasta, I. R. Cantalapiedra, C. E. Auguet, A. Peñaranda, and L. Ramírez-Piscina. Modeling of spatiotemporal patterns in bacterial colonies. *Phys. Rev. E*, 59(6):7036–7041, 1999.
42. I. Cohen. Mathematical modeling and analysis of pattern formation and colonial organization in bacterial colonies, 1997. M.Sc. thesis, Tel-Aviv University, ISRAEL.
43. J. D. Murray. *Mathematical Biology*. Springer-Verlag, Berlin, 1989.
44. J. A. Shapiro and D. Trubatch. Sequential events in bacterial colony morphogenesis. *Physica D*, 49:214–223, 1991.

45. A. Grondin, H. C. Jarell, and L. R. Berube. Discontinuous expansion linked to sector formation in *Pseudomonas aeruginosa* colonies. *Archives of Microbiology*, 172:59–62, 1999.
46. I. Golding, I. Cohen, and E. Ben-Jacob. Studies of sector formation in expanding bacterial colonies. *Europhys. Lett.*, 48(5):587–593, 1999.
47. J. Cairns, J. Overbaugh, and S. Miller. The origin of mutants. *Nature*, 335:142–145, 1988.
48. E. Ben-Jacob. From snowflake formation to the growth of bacterial colonies. part I: Diffusive patterning in non-living systems. *Contemp. Phys.*, 34:247–273, 1993.
49. M. Eisenbach. Functions of the flagellar modes of rotation in bacterial motility and chemotaxis. *Molec. Microbiol.*, 4(2):161–167, 1990.
50. J. B. Stock, A. M. Stock, and M. Mottonen. Signal transduction in bacteria. *Nature*, 344:395–400, 1990.
51. C. H. Shaw. Swimming against the tide: chemotaxis in *agrobacterium*. *BioEssays*, 13(1):25–29, 1991.

Periodicity in Age-Resolved Populations

Sergei Esipov

CRA, One Chase Manhattan Plaza New York, NY 10005, USA

Abstract. We discuss the interplay between the non-linear diffusion and age-resolved population dynamics. Depending on the age properties of collective migration the system may exhibit continuous joint expansion of all ages or continuous expansion with age segregation. Between these two obvious limiting regimes there is an interesting window of periodic expansion, which has been previously used by us in modeling bacterial colonies of Proteus mirabilis. In order to test whether the age-dependent collective migration leads to periodicity in other systems we performed a Fourier analysis of historical data on ethnic expansions and found multiple co-existing periods of activity.

To my Father

1 Introduction

The first explanation of spatial instability exhibited by a dissipative system belongs to Turing [1]. His remarkable article contains a suggestion that the steady-state spatial patterns can arise simply from reaction and diffusion under certain conditions. The impact of the Turing's idea on mathematical and physical biology was profound [2,3], and it continues to be the "number one choice" when a model is sought for a spatially periodic pattern. Back in 1952 there was little novel in discovering a linear instability within a positive band of wave numbers coexisting with stability at the zero wave number: such systems had been known in hydrodynamics due to the works by Taylor [4]. Yet, the source of this instability was unusual: it was the diffusivity mismatch between (at least two) different reacting species.

When bacterial colonies were selected by specialists working on pattern formation as the next application of ideas borrowed from critical phenomena and crystal growth, a "dictionary" was formed based on previously described patterns to facilitate modeling. In this dictionary the periodic patterns were associated with the Turing-like reaction-diffusion models. Recently, such a model was applied [5] to the ziggurats built by the enterobacteriaceae called *Proteus mirabilis* [6]. The purpose of the present article is to address our previous conclusion that Proteus periodicity finds a much more detailed explanation in age-resolved dynamics [13], and to present a minimal age-resolved model. We also demonstrate that a similar type of behavior could be found in other, much more complex systems.

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 242–251, 2001.
© Springer-Verlag Berlin Heidelberg 2001

Proteus colonies on hard agar are built by repeating swarming and consolidation phases. Their periodicity has one distinct feature, first discussed by Shapiro: the environmental conditions, nutrients, agar concentration, widths of terraces might vary[1], but the timing between terraces is preserved [6]. From the microbiological point of view, if one cannot find simple convincing arguments for the observed effects, it is usually assumed that the system has an additional internal mechanism, in this case, a certain dedicated bacterial clock, which is responsible for ziggurats. This is when a physicist may sometimes offer help by providing an alternative "mechanistic" explanation. The following brief description is based on the results published in [6,13].

It is known that the *Proteus* colony on agar consists of cells of two types, called swimmers (which are not motile on agar, they can only "swim" in non-viscous liquid) and poly-nucleic swarmers. It is also known that swarmers, as their name suggests, are capable of migrating over the surface, and they do it in groups, called rafts. When swarmers manage to occupy a novel region of the agar surface, they tend to stop, dedifferentiate, form swimmers, and thus initiate what is termed as consolidation phase. In the course of consolidation phase there is no migration, but, eventually, new swimmers and swarmers are formed, swarmers, again, organize themselves in rafts and the next swarming phase begins. Both phases together have a well-defined period, at 32°C it takes about 4.7 hours to complete one cycle, see Fig. 1.

This period is comparable with the swimmer cell division time (1.5 - 2 hours) and swarmer longevity (a variable, it is few times larger than swimmer division time, as swarmers containing, say, 16 nuclei are not uncommon, while it should take at least four divisions to form them). Neither of these division times is a well-defined constant, and it was unclear what might organize divisions of different cells and lead to a coherent periodic expansion, as this hypothetical clock must be inherited by daughter cells without aberrations. Thus, a model was required which would incorporate cell divisions, the only known processes at similar time scales, and somehow synchronize them. It was found by us previously that the non-linear diffusion was the synchronizer: only mature swarmers are able to associate in rafts and when migration stops, and cells septate, their ages are made equal. We will not re-introduce the entire model here, details can be found in [7], instead, it is our intention to select the key important ingredients of the age-resolved kinetic model and discuss them. This can be found in Section 2. Section 3 contains quantitative analysis of historical data on ethnic military expansions, which also exhibit a certain periodicity, and where age structure of the population is important.

2 The Age-Resolved Model with Diffusion

Let us consider a distribution function for population groups, $n(x, \theta, t)$, of age θ, which may have a separate dynamics in space, x, and time, t. The evolution

[1] Ambient temperature represents an exception.

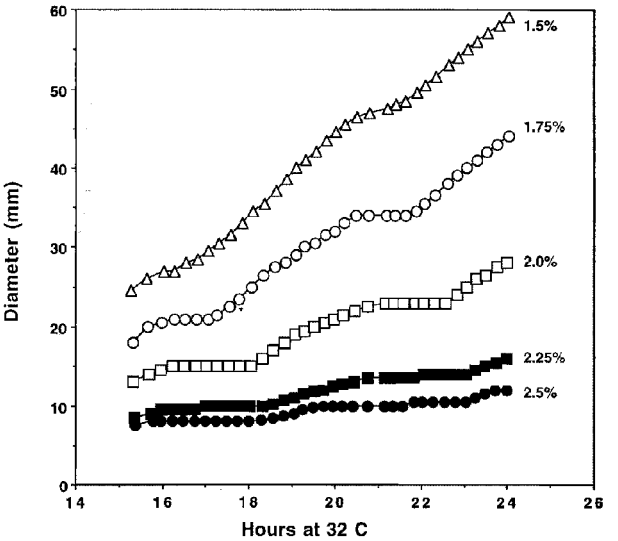

Fig. 1. Time dependence of the Proteus mirabilis colony size on Petri dishes with different concentration of agar, shown in percents [6]. Larger concentrations of agar correspond to slower motility of swarmer cells. Despite significant changes in the durations of swarming and consolidation phases, their sum, i.e. the period, is clearly preserved.

equation reads, simply,

$$\partial_t n + \partial_\theta n = \partial_x (D\, \partial_x n) + F[n]\,, \tag{1}$$

where the term may contain age migration, insertion, extraction, and "reactions". Unlike the purely hydrodynamic analogies of reaction-diffusion systems, Eq. (1) is similar to what one can find in physical kinetics, i.e. to the Boltzmann equation. We now discuss a few cases, where analysis is straightforward, by using a periodic θ-axis, $0 \le \theta \le \Theta$, and age-resolved diffusivity, $D(\theta)$. Since periodic θ-axis is a built-in oscillator, $D(\theta)$ already leads to a simple system with periodic diffusivity, but not to a periodic expansion. In the absence of "reaction" term, $F[n]$, an inhomogeneous initial condition will diffuse with the diffusion length given by

$$L^2 = \int_0^t D(\theta - t')\, dt'\,. \tag{2}$$

Since time, t, and periodic age, θ, both have the dimensionality of time, we adopt the convention that the argument $\theta - t$ samples diffusivity at the age $\mathrm{mod}(\theta - t, \Theta)$. The final state in this linear case is spatially uniform while the initial (spatially averaged) age-structure is preserved. Diffusion aids to relaxation here.

The situation somewhat changes with introducing a non-linear diffusion. The straightforward application of the popular model [3] reads as follows

$$D_1(n) = D_0[n(x,\theta,t)/n_0]^m \,. \tag{3}$$

where m is a positive number, D_0, n_0, and other values below with the subscript 0 are parameters. In order to keep the notations informative, we will sometimes suppress the notations of dependencies, such as the dependence of D on x, θ, t through n in (3). We could have also selected *c*ollective non-linear diffusivity by using a functional dependence ([...])

$$D_2[n] = D_0 \left[\frac{1}{N_0} \oint d\theta \, n(\theta) \right]^m . \tag{4}$$

With the diffusivity Eq. (3), all ages are decoupled. Each generates a moving hull with the well-defined front, $|x| \leq v\tau$,

$$\begin{aligned} n(x,\theta,t) &= \frac{n_0}{\tau} \left[1 - \left(\frac{x}{vt} \right) \right]^{1/m} , \\ \tau^{m+2} &= \frac{t}{t_0} \,, \\ t_0 &= \frac{mv^2}{2(m+2)D_0} \,, \\ v &= \frac{N(\theta - t)}{a \, n_0} \,. \end{aligned} \tag{5}$$

Here $N(\theta) \equiv \int dx \, n(x,\theta,0)$ is the initial population broken by age, and $a = B(\frac{1}{2}, \frac{m+1}{m})$ is expressed through the beta-function. Obviously, this dynamical process leads to age segregation. The age group with largest total concentration will be on the leading front.

Unlike Eq. (3), the diffusivity Eq. (4) dictates that all age groups have the same mobility. Integrating (1) over ages one recovers that the total concentration again follows the dynamics in Eq. (5) but with $\mathbf{N} \equiv \oint d\theta n(x,\theta,0)$ being age-independent, and with n_0 being replaced by N_0.

$$\begin{aligned} N(x,t) &= \frac{N_0}{\tau} \left[1 - \left(\frac{x}{vt} \right) \right]^{1/m} , \\ \tau^{m+2} &= \frac{t}{t_0} \,, \\ t_0 &= \frac{mv^2}{2(m+2)D_0} \,, \\ v &= \frac{1}{a \, N_0} \oint d\theta \, N(\theta) \,. \end{aligned} \tag{6}$$

Every point in space has identical age structure, distributed in proportion to the initial spatially integrated total concentration $N(\theta)$.

Integration over ages does not lead to a closed equation when explicit θ-dependence is added to Eq. (4), e.g.

$$D_3[n] = D_0\, H(\theta - \theta_0) \left[\frac{\mathbf{N}}{N_0}\right]^m . \tag{7}$$

where $H(x)$ is the Heaviside function and Θ is the age period. The model Eq. (7) is appropriate in systems with moving fronts (hence, non-linear diffusivity) where only "mature" age groups can participate in collective migration. Qualitatively, at large non-linearity (large m), and at a given time, the age groups from θ_0 to Θ form a hull-shaped moving front, as in Eq. (5), while they continue to loose "senior" groups at age Θ and accumulate "junior" groups at age θ_0. Both processes are taking place adiabatically. Upon passing through the maximal age, Θ, the age groups become immobile, and their spatial profile is frozen during the following θ_0 units of time. The junior groups past age θ_0 share the strong non-linear diffusivity and quickly occupy the entire space inside the moving front(s). Since the age memory is preserved in the models considered above, the motion of the front reflects the initial age structure, i.e. may exhibit periodicity, provided the age structure is not uniform. On top of this, the spatial transport of each group selected by age memory, $n(\theta - t)$, is also periodic, simply reflecting the time windows of zero diffusivity at $\mathrm{mod}(t + \theta, \Theta) < \theta_0$.

Although we haven't covered any significant part of possible functional forms for non-linear age-resolved diffusivity, one finds that the intermediate case Eq. (7), interpolating between fully age-resolved Eq. (3) and age-independent Eq. (4) cases leads to age segregation followed by unification, i.e. to periodicity. It is now of interest to understand - what part of this phenomenology survives when we partially relax the complete coherence of the age dynamics. The simplest way to do it mathematically is to invoke some age "diffusion", i.e. add certain randomness to the deterministic drift along the age axis. To avoid unnecessary arguments about violating the causality principle, we assume that the term $F[n]$ contains jumps, which transfer an individual from any age θ to a more *senior* age θ', bypassing the interval (θ, θ'). At large times the net (Fokker-Planck) effect of such jumps is clearly an increase in aging rate (previously assumed to be 1) plus an age diffusion. While the change of average aging rate can be suppressed by rescaling the time axis, the age-diffusion is the phenomenon we would like to study in more details. This leads to the equation

$$\partial_t n + \partial_\theta n = \partial_x(D_3\, \partial_x n) + \partial_\theta(d\, \partial_\theta n)\,, \tag{8}$$

where $d \ll \Theta$ is the small age diffusivity. One would like to find out whether the above-mentioned periodicity would survive under the influence of the age diffusivity as the age memory is no longer preserved, and all age groups tend to merge in the absence of spatial diffusion, so that the integral $\int_{\theta_0}^{\Theta} d\theta\, n(\theta, t)$ ceases to depend on time, and the initial condition is forgotten. As a result, the periodicity in spatial migration would eventually relax (although stochastic age groups formed by age memory may continue to enjoy zero diffusivity periods and related quasi-periodicity).

We are now in a position to formulate the minimal model of the type corresponding to our *Proteus* model in [13]. What one has to do to ensure periodicity is to add cell division and suppress diffusion at high enough concentration of cells. Then, the diffusivity takes the form

$$D_4[n] = D_0\, H(\theta - \theta_0)\, [y(1-y)]^m, \qquad (9)$$
$$y \equiv \frac{1}{N_0} \int_{\theta_0}^{\Theta} d\theta\, n(\theta)\,,$$

and the equation, finally, reads,

$$\partial_t n + \partial_\theta n = \partial_x (D_4\, \partial_x n) + \partial_\theta (d\, \partial_\theta n)\,, \qquad (10)$$

In an uniform system the shut-down concentration N_0 is reached by the time $\frac{1}{\gamma} \ln[N_0/N(0)]$. As long as we have

$$\frac{d}{\gamma} \ln \frac{N_0}{N(0)} \ll \Theta^2\,, \qquad (11)$$

the age structure remains non-uniform, and expansion may exhibit periodicity, although its initial condition would be eventually forgotten. This ends our discussion of the minimal age-resolved model. Analysis of Eqs. (9,10) will be published elsewhere. The next Section is devoted to testing the hypothesis that even most advanced age-resolved collective migrations may exhibit simple periodicity!

3 Ethnic Expansions 500BC - 1400AD

The military activity in the Mediterranean region, and, later, in Europe, Western Asia, and Northern Africa is documented very extensively, and one can find reliable data describing events almost year by year. The historical tables provide sufficient information for a statistical study. Certainly, all the properties needed above to form the age-resolved spatial dynamics are dwarfed by complexity of the human ethnic migration. It is convincing that there exists the minimal migration age θ_0, when a warrior could carry his own weapon, while scars and rewards of military campaigns undoubtedly contributed to rapid aging and age "diffusion". Military units usually prefer to move in groups, and in large enough numbers, hence the non-linear, collective, age-dependent diffusivity is surely a part of the story. There are many details, which are arguably very important, but haven't been addressed above: this includes sexual segregation, social structures of all involved ethnic groups (e.g. Mongols offered any potential enemy to join the Golden Horde before attacking their settlements), separate dynamics for the defeated (in the model above the species were expanding into unoccupied territory), just to name a few. Still, it was of interest to process the available data by computer to test for the periodicity in ethnic migration.

With this in mind we selected a book by Grun [8], which offers data on history and politics along with many other activities broken by year. Data were treated

as follows. If in a particular year a certain ethnic groups are mentioned as either attacking neighboring territory or successfully defending their own territory, this year is marked as +1 in the table for a successful group, otherwise it is assigned number zero[2]. We screened a total of 1901 years from 500BC to 1400AD. Only ethnic groups that had the cumulative number of successful military activities in excess of 21 were used for subsequent analysis (the number 21 was selected as sufficient amount of data for the subsequent Fourier analysis). Other ethnic groups and civilizations for some reason didn't make it to the Grun and Stein tables, although we deal with a very good and unusually quantitative history book.

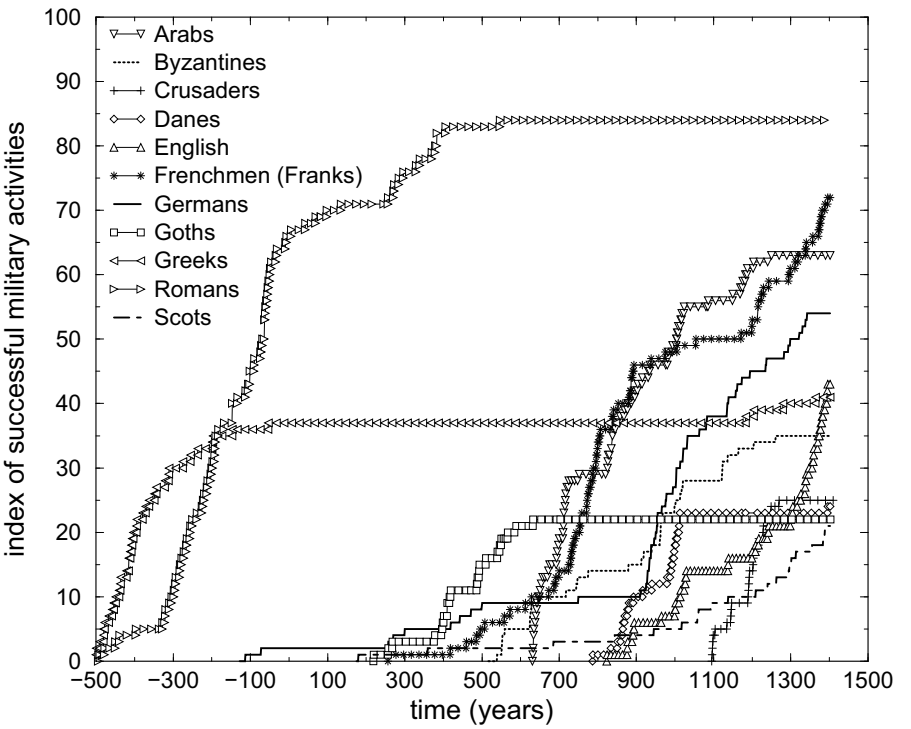

Fig. 2. Time dependence of the cumulative number of successful military activities of all the ethnic groups frequently mentioned in [8].

[2] In cases of continuous wars lasting for more than one year we used only years when it was explicitly mentioned that a certain ethnic group was either attacking neighboring territory or successfully defending its own territory. Thus, various multi-year wars such as Punic wars may be represented by only a few entries. The only exception was made for Crusades. As a test, these entries were subject to a special treatment: we entered +1 for every year that a given Crusade lasted, as indicated in the book. The resulting periodicity is not significantly different from the results for other entries.

The curves for the cumulative numbers of successful military activities can be seen in Fig. 2. We think that there is a suggestive similarity between Fig. 1 and Fig. 2. One can go further and extract the power spectrum of the curves in Fig. 2 by using, say, the discrete Fourier transform. An example of such spectrum is shown in Fig. 3.

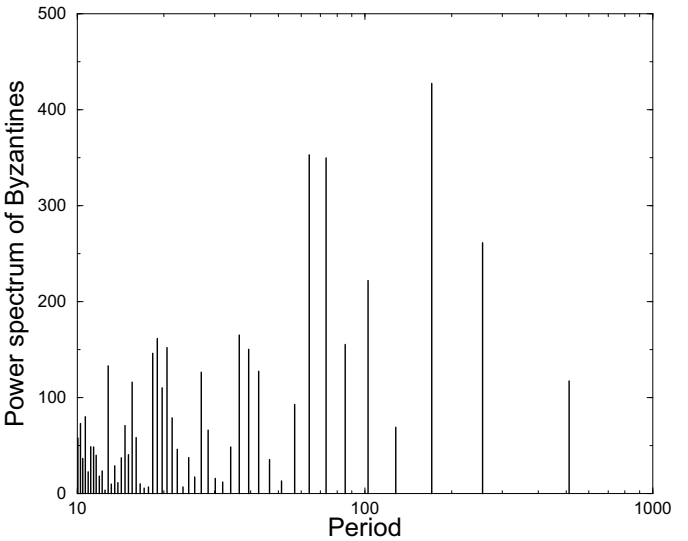

Fig. 3. Power spectrum of the curve in Fig. 2 for Byzantines vs. period (inferred from frequency, in years.

Finally, Table 1 gives compilation of all the well-expressed periods extracted from the power spectrum. E.g. the two most pronounced peaks in Fig. 3 correspond to periods 70 and 170 yrs. One can see that multiple periods for a given ethnic group are omnipresent. Their origin and historical significance are outside of the scope of this article. We only mention that the periods in the left half of Table 1 are similar to the human reproductive age, and typical lifespan.

Table 1 raises more questions than it can possibly answer. The origin of collective migration (such as the invention of Macedonian phalanx or formation of common military goals based on religious beliefs) is left unspecified. The same applies to the influence of philosophy, learning, medicine, science, technology, even literature, etc. Bursts of ethnic activity have been recognized by historians, such as Bordieu, Leroi-Gourhan, Levi-Strauss, and in the Gumilev's passionate theory of ethnogenesis [9]. We believe that the present approach provides a complimentary technical analysis. It is also of interest to study spatial patterns associated with ethnic migrations.

It has to be kept in mind that our objective was to perform a *quantitative* analysis on periodicity of historical ethnic migrations and compare them with

Table 1. Periods contributing the largest signals in the power spectra.

Fourier years		10-20	20-40	40-60	60-100	100-150	150-210
512	Arabs				70		170
512	Byzantynes				70		170
256	Crusaders	12		51			
256	Danes					128	
512	Englishmen	13				102	
1024	Franks/Frenchmen	14	20	54		146	
1024	Germans		26				
512	Goths				73		
512	Greeks	11	21				
1024	Romans	14		48		113	208
512	Scots		20				

typical time scales of human life. This is an important step is modeling periodic expansions by means of kinetic models, beyond the reaction-diffusion schemes.

The present work originated as an attempt to formulate a mathematical model for the concept of the "geometry of biological time", introduced by Winfree [10]. The model is different from a system of coupled non-linear oscillators, our "θ-oscillators" are, in fact, linear, fuzzy, and can diffuse in space in synchronized groups, as given by Eq. (9).

Acknowledgments

I am thankful to Clara Salueña for invaluable help with the manuscript, multiple discussions and encouragement.

Opinions expressed in this article are the author's and do not necessarily correspond to the views of CRA. All errors in this text are the author's responsibility.

References

1. Turing, A. M.: The Chemical Basis of Morphogenesis. Phil Trans.Roy. Soc. Lond. **B237**, 37-72 (1952).
2. Levin S. A., and Segel L. A.: Pattern Generation in Space and Aspect. SIAM Rev. **27**, 45-67 (1985).
3. Murray, J. D.: *M*athematical Biology. Springer-Verlag, Berlin (1993).

4. Taylor, G. I.: Stability of a viscous liquid contained between two rotating cylinders, Phil. Trans. R. Soc. Lond. **A223**, 289-309 (1923).
5. Lacasta A. M., Cantalapiedra I. R., Auguet C. E., Peñaranda A., and Ramírez-Piscina L.: "Concentric Rings Modeling in Bacterial Colonies". Poster A-10, XVII Sitges Euroconference, Sitges, Barcelona, Spain, 5-9 June 2000.
6. Rauprich, O., Matsushita, M., Weijer, K., Siegert, F., Esipov, S., and Shapiro, J. A.: Periodic phenomena in Proteus mirabilis swarm colony development. J. Bacteriol. **178**, 6525-6538 (1996).
7. Esipov, S. E., and Shapiro, J. A.: Kinetic Model of Proteus mirabilis Swarm Colony Development, J. Math. Biol. **36**, 249-268 (1996).
8. Grun, B.: *The Timetables of History. A Horizontal Linkage of People and Events. Based on W. Stein's Kulturfahrplan*, Simon and Schuster, New York (1982).
9. Gumilev, L. N.: *Ethnogenesis and the biosphere*, Moscow, Progress, (1990).
10. Winfree, A. T.: *The Biology of Biological Time*, Springer-Verlag, New York, (1990).

Nonequilibrium Nanostructures in Condensed Reactive Systems

Alexander S. Mikhailov, Michael Hildebrand, and Gerhard Ertl

Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

Abstract. Chemical reactions in condensed systems, where attractive interactions between particles are present, may lead to the development of nonequilibrium nanostructures. Our theoretical investigations of surface chemical reactions show the existence of stationary and traveling nanostructures in these systems. Nonequilibrium localized structures, representing chemical nanoreactors, have also been found.

1 Introduction

Reaction-diffusion systems are known to show a great variety of nonequilibrium spatiotemporal patterns - stationary dissipative structures, traveling waves, target patterns, rotating spirals, and wave turbulence [1,2]. The common property of such classical systems is that molecules perform free random diffusive motion between occasional reactive collisions. Hence, they can be compared with reactive ideal gases, where the mean free path of atoms sets a limit for the minimum spatial inhomogeneity. In the case of reaction-diffusion systems, the smallest characteristic dimension of the patterns is set by the diffusion length, which is simply the mean distance passed by a molecule until it undergoes a reaction. The diffusion length is relatively large: it ranges from fractions of a millimeter for the Belousov-Zhabotinsky reaction to about a micrometer in typical biochemical enzymatic reactions [3]. Thus, nanoscale self-organization cannot be realized in systems where only diffusion and reactions are present.

In contrast to this, structures in condensed systems with attractive physical interactions between particles can be very small. Atomic clusters may include just tens or hundreds of atoms. "Quantum dots" on specially prepared metal surfaces have a characteristic size of a few nanometers. Generally, the minimum limit for the size of equilibrium structures is determined by the radius of attractive interactions between molecules in such systems, which can be a fraction of a nanometer.

To obtain nonequilibrium nanoscale structures, systems with both chemical reactions and physical interactions between the particles must be considered. These structures may result from the combined action of cohesion, reactions and diffusion. The cohesion, responsible for the formation of equilibrium thermodynamics structures, should however not be too strong. Since the reaction effects are typically relatively weak, they can compete with the physical interactions between the particles and thus influence the microscopic organization of a system only if it is structurally labile.

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 252–269, 2001.
© Springer-Verlag Berlin Heidelberg 2001

Structural lability is a feature of "soft matter" representing molecular systems with weak attractive interactions. Examples of such soft matter are provided by liquid crystals, Langmuir-Blodgett films, lipid membranes, vesicles, or polymers. These condensed systems often have a great number of different equilibrium structures and can easily undergo phase transitions when the medium parameters are changed. Stable nonequilibrium structures emerge from an interplay between such structural phase transitions and chemical reactions, coupled to diffusion [4].

Nonequilibrium nanostructures should be important in biology, at the level of individual living cells. The typical size of a cell is just a few micrometers which is comparable with the diffusion length for intracellular biochemical reactions. Indeed, pure reaction-diffusion patterns which were so far observed in individual cells have relatively large spatial dimensions. Calcium spiral waves with the wavelength of tens of micrometers were observed in experiments with very large cells, i.e. in frog eggs with a diameter of a millimeter [5]. Recently, traveling glycolic waves with a wavelength of several micrometers were found in smaller cells of size of about thirty micrometers [6]. Though such reaction-diffusion patterns may have some biological function, they cannot form the basis of the microscopic self-organization of intracellular processes. The most relevant processes of nonequilibrium pattern formation in single biological cells must proceed on submicrometer and nanometer scales. With respect to this aspect, it is especially remarkable that the cell is full of various kinds of soft matter interacting with complex chemical reaction networks.

On the other hand, artificial nonequilibrium nanostructures may find interesting technological applications. As compared with the nanostructures which result from equilibrium microphase separation (see [7]), they may show a more broad variety of spatial patterns that can be easily controlled by changing the reaction conditions. Hence, they can be used to produce a greater range of microstructured materials. Additionally, and perhaps even much more important, the nonequilibrium nanostructures are not necessarily stationary, as all equilibrium structures should be. Instead, such nanostructures can oscillate, show complex motions and even steadily travel through the medium. Thus, they can be used to transport materials on microscales and to operate as micromachines.

Despite the apparent importance of nonequilibrium structures in reactive soft matter, up to date they have been investigated very little, both theoretically and experimentally. In contrast to reaction-diffusion systems, where an extensive general theory is already available, only several examples of such nonequilibrium structures have been studied so far. Stationary reaction-induced micropatterns have been theoretically predicted [8,9] and experimentally observed [10] in polymer blends undergoing phase separation. Traveling waves have been observed in an illuminated Langmuir-Blodgett film [11], though a satisfactory theory of this effect is not yet available. In both these cases, the characteristic wavelength of a pattern was still relatively large, lying in the range of tens of micrometers. Traveling "atomic strings" with the sizes of less than a micrometer have however recently been observed using atomic force microscopy (AFM) on crystal surfaces under reaction conditions [12].

The experimental observation of nanoscale nonequilibrium structures meets certain difficulties. Since they are so small, powerful microscopes should be used. However, because interactions responsible for such structures are weak and the structures are labile, they may be easily influenced and even destroyed by the observation process. Moreover, the lability of nonequilibrium nanostructures makes them also strongly affectable by thermal fluctuations and by the stochastic reaction noise. Therefore, the employed microscope should have sufficient temporal resolution, so that the fluctuations do not smear the observed pattern.

Systems well suited for the experimental investigations of nonequilibrium nanoscale pattern formation are provided by chemical reactions in adsorbate monolayers on metal surfaces. Because the reaction patterns are strictly two-dimensional, the reaction process in this system can be fully monitored. Scanning tunneling microscopy (STM) allows to observe the reaction patterns, displaying the actual positions of individual reacting atoms. The standard STM equipment has however a relatively poor temporal resolution, so that only "frozen" atomic distributions could be recorded previously. Recently, significant progress in the development of a fast STM has been achieved. As a result, the observation of surface chemical reactions with an atomic resolution became possible in real time. The experiments show that the reactions are typically accompanied by the appearance of nanoscale structures [13,14,15,16,17].

Surface chemical reactions are important because they form the basis for heterogeneous catalysis, broadly used in chemical industry. But, on the other hand, they can also be viewed as model systems whose theoretical and experimental investigation may provide insights into general aspects of nonequilibrium pattern formation on nanoscales, similar to the role played by the classical Belousov-Zhabotinsky reaction in the studies of macroscopic pattern formation in reaction-diffusion systems. The aim of this article is to present a brief review of our recent theoretical investigations of such reactive condensed systems.

2 Phase Transitions and Chemical Reactions in Adsorbed Monolayers

Molecules or atoms may adsorb on single crystal surfaces of metals up to a monolayer. At sufficiently high temperature, adsorbed particles can freely move over the crystalline surface and thus a two-dimensional gas (or liquid) of adsorbed particles is formed. In addition to strong interactions between these particles and the metal substrate, much weaker lateral interactions between the particles are also present in most cases. Typically, such interactions are substrate-mediated: An adsorbed particle modifies the local electronic structure in the surface of the substrate changing the surface potential for other adsorbed particles in its vicinity. Effective interactions between adsorbed particles may also result from elastic deformations induced by them in the substrate. The substrate-mediated lateral interactions can extend over distances ranging from a few lattice lengths to hundreds of nanometers, when elastic forces are involved. These interactions can be repulsive or attractive.

If weak attractive lateral interactions between adsorbed particles are present, the two-dimensional gas of such particles can undergo a phase transition into a condensed liquid-like state if the temperature is decreased. This first-order phase transition proceeds through critical nucleation or spinodal decomposition, depending on the parameters of the system. When desorption is absent and the number of adsorbed particles is conserved, separation of phases takes place in the monolayer.

Another kind of phase transition results from the action of adsorbed particles on the crystal substrate. When the surface is uncovered, the energetic interactions between metal atoms forming the top layer of the substrate are different from those in the bulk. Therefore, under certain conditions these metal atoms would rearrange and form a two-dimensional array which is different from the crystalline atomic structure in the bulk of the metal. Such a "reconstructed" surface structure can be lifted by the presence of adsorbed particles [18]. This latter effect can correspond to a first-order phase transition.

When two kinds of molecules are adsorbed on the surface, chemical reactions between them are possible. Since the adsorption significantly modifies the properties of these species, surface chemical reactions can be greatly different from the respective reactions between individual molecules in the gas. Even if a reaction in the gas phase is not possible, because the respective energy barrier is too high, it may easily proceed when two pre-adsorbed particles meet on the surface. In this case the metal is a *catalyst* facilitating a particular chemical reaction.

Our attention in this short review paper will be focused on situations where an interplay between surface chemical reactions and phase transitions is involved. As will be shown, such an interplay can lead to the development of various stationary or traveling nonequilibrium nanostructures in this type of systems.

Mathematical modeling of surface chemical reactions on nanoscales can be performed in the framework of kinetic Monte Carlo simulations on lattices. Though such simulations yield valuable information about the microscopic course of chemical reactions, they also have a certain disadvantage. The Monte Carlo simulations are essentially numerical experiments and, as such, they do not constitute an analytic theory, i.e. they do not allow to make predictions concerning the properties and the evolution of structures.

We have developed an approximate mesoscopic approach to model such surface chemical reactions which is formulated in terms of continuous fluctuating concentrations (coverages) of adsorbed particles [19,20]. To derive the stochastic evolution equation for these variables, we start with the full master equation for the joint probability distribution of particles that are diffusively hopping over the lattice and can react, adsorb or desorb (the hopping rates can be influenced by interactions between the particles). The coarse graining is performed by introducing a system of boxes, each containing a large number of lattice sites but still smaller than the radius of lateral interactions between the particles. The fluctuating local concentrations (coverages) are then defined as the fraction of lattice sites inside a box occupied by the considered kind of particles. Using the Kramers-Moyal expansion, the master equation is approximately trans-

formed into a functional Fokker-Planck equation which, in turn, corresponds to a stochastic partial differential equations for the fluctuating local coverages.

Though the derivation is technically complicated, the final result is simple. If there is only one kind of particles which can adsorb, desorb, diffuse, react and attract each other, the stochastic evolution equation for the local coverage $c(\mathbf{r},t)$ is (see [20])

$$\frac{\partial c}{\partial t} = k_a p_0(1-c) - k_d c - k_r c + \frac{\partial}{\partial \mathbf{r}}\left(\frac{D}{k_B T}(1-c)c\frac{\partial U}{\partial \mathbf{r}}\right) + D\frac{\partial^2 c}{\partial \mathbf{r}^2} + \xi(\mathbf{r},t) \quad (1)$$

The first term on the right side of this equation describes adsorption of particles (p_0 is the partial pressure in the gas phase above the crystal, k_a is the sticking coefficient), the second term corresponds to desorption of particles and the third term describes the reaction (see the comment below). The next term takes into account the *viscous flow* of particles over the surface induced by the gradient of the local potential $U(\mathbf{r})$ for adsorbed particles. This nonuniform potential is due to attractive lateral interactions between the particles,

$$U(\mathbf{r}) = -\int u(\mathbf{r}-\mathbf{r}')c(\mathbf{r}')d\mathbf{r}' \quad (2)$$

where $u(\mathbf{r})$ is the two-particle attraction potential,

$$u(r) = \frac{u_0}{\pi r_0^2}\exp\left(-\frac{r^2}{r_0^2}\right) \quad (3)$$

with the parameter u_0 specifying the interaction intensity and the parameter r_0 determining the interaction radius; T is the temperature and k_B is the Boltzmann constant. The desorption rate depends on the local potential as

$$k_d = k_{d,0}\exp\left[\frac{U(\mathbf{r})}{k_B T}\right] \quad (4)$$

Generally, the reaction rate constant k_r of a monomolecular reaction should also depend on the local potential U. We assume that the reaction is nonequilibrium (e.g. photo-induced) and therefore such a dependence is absent.

The differential equation (1) includes the internal noise $\xi(\mathbf{r},t)$. The noise takes into account fluctuations resulting from the stochastic nature of the system, approximately described by this evolution equation. Its intensity and functional form are determined through the master equation specifying the detailed stochastic evolution of the system. In the considered case we have (see [20])

$$\begin{aligned}\xi(\mathbf{r},t) = {} & \frac{1}{Z^{1/2}}\sqrt{k_a p_0(1-c)}f_a(\mathbf{r},t) + \frac{1}{Z^{1/2}}\sqrt{k_d}f_d(\mathbf{r},t) \\ & + \frac{1}{Z^{1/2}}\sqrt{k_r c}f_r(\mathbf{r},t) + \frac{1}{Z^{1/2}}\frac{\partial}{\partial \mathbf{r}}\left(\sqrt{2Dc(1-c)}\mathbf{f}(\mathbf{r},t)\right)\end{aligned} \quad (5)$$

Each process (i.e. adsorption, desorption, reaction and diffusion) gives rise to its own internal independent noise. In this equation $f_a(\mathbf{r},t)$, $f_d(\mathbf{r},t)$, and $f_r(\mathbf{r},t)$

are white random forces of unit intensity, such that $< f_a(\mathbf{r},t)f_a(\mathbf{r}',t') >=< f_d(\mathbf{r},t)f_d(\mathbf{r}',t') > = < f_r(\mathbf{r},t)f_r(\mathbf{r}',t') >= \delta(\mathbf{r}-\mathbf{r}')\delta(t-t')$. The noise of diffusion results from random fluxes; the respective vectorial white random force $\mathbf{f}(\mathbf{r},t) = \{f_x(\mathbf{r},t), f_y(\mathbf{r},t)\}$ has $< f_x(\mathbf{r},t)f_x(\mathbf{r}',t') >=< f_y(\mathbf{r},t)f_y(\mathbf{r}',t') >= \delta(\mathbf{r}-\mathbf{r}')\delta(t-t')$ and $< f_x f_y >= 0$. The parameter Z is the number of lattice sites per unit surface area, it is related as $Z = 1/l_0^2$ to the lattice length l_0.

As we have already noted above, the stochastic evolution equation is obtained by applying coarse-graining over surface elements of a certain linear size l, such that $r_0 \gg l \gg l_0$. Therefore, it correctly retains only fluctuations with the characteristic lengths exceeding l and the characteristic times exceeding the mixing time $\tau = l^2/D$ within the surface element of length l. Note that all internal noises are multiplicative.

In the absence of the reaction ($k_r = 0$) equation (1) describes an equilibrium first-order phase transition in the two-dimensional non-ideal gas of adsorbed particles. When desorption and adsorption of particles are absent, so that their total number is conserved, this equation takes the simple form

$$\frac{\partial c}{\partial t} = -\frac{\partial}{\partial \mathbf{r}}\left(\frac{D}{k_B T}(1-c)c\frac{\partial}{\partial \mathbf{r}}\int u(\mathbf{r}-\mathbf{r}')c(\mathbf{r}')d\mathbf{r}'\right) + D\frac{\partial^2 c}{\partial \mathbf{r}^2} + \frac{1}{Z^{1/2}}\frac{\partial}{\partial \mathbf{r}}\left(\sqrt{2Dc(1-c)}\mathbf{f}(\mathbf{r},t)\right) \quad (6)$$

Because of the presence of an integral term, this differential equation is nonlocal. It can be shown that it can be reduced to the local Cahn-Hilliard equation, used in the kinetic theory of first-order phase transitions [21], in the vicinity of the critical point, when fluctuations are neglected and the formal limit $r_0 \to 0$ is taken (see [20],[22]).

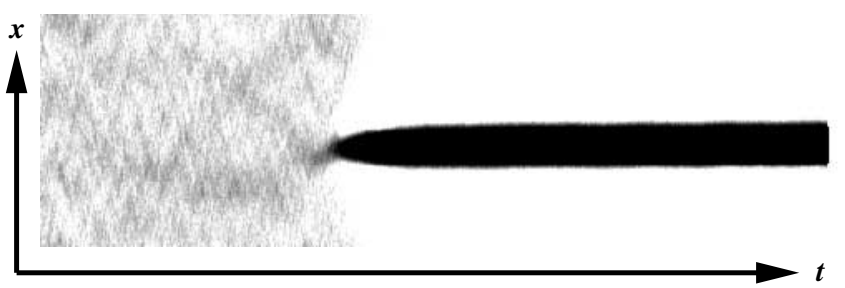

Fig. 1. Spatiotemporal evolution of the coverage according to equation (6) in a one-dimensional system of size $L = 50$ for $u_0/k_B T = 4.1$, $r_0 = 0.8$, $D = 1$, $Z = 1000$ and average coverage $< c >= 0.3$. The local surface coverage c is shown in gray scale, with darker areas corresponding to higher coverages

Our numerical simulations of equation (6) show that it indeed describes stochastic nucleation, spinodal decomposition and Ostwald ripening in the considered system. Fig. 1 displays critical fluctuations ending with the appearance

of a supercritical nucleus and phase separation in a one-dimensional system [23]. We see that fluctuations become depressed after the phase separation has taken place.

Fig. 2 [23] shows examples of the spatiotemporal stochastic evolution of a two-dimensional system for two different average coverages $< c >$. When $< c >= 0.3$ (Fig. 2a), stochastic nucleation of the dense phase on the background of the dilute phase occurs. Subsequently, competition between the growing supercritical nuclei sets in, so that the smaller nuclei disappear. When $< c >= 0.5$ (Fig. 2b), the uniform phase is absolutely unstable and spinodal decomposition is observed.

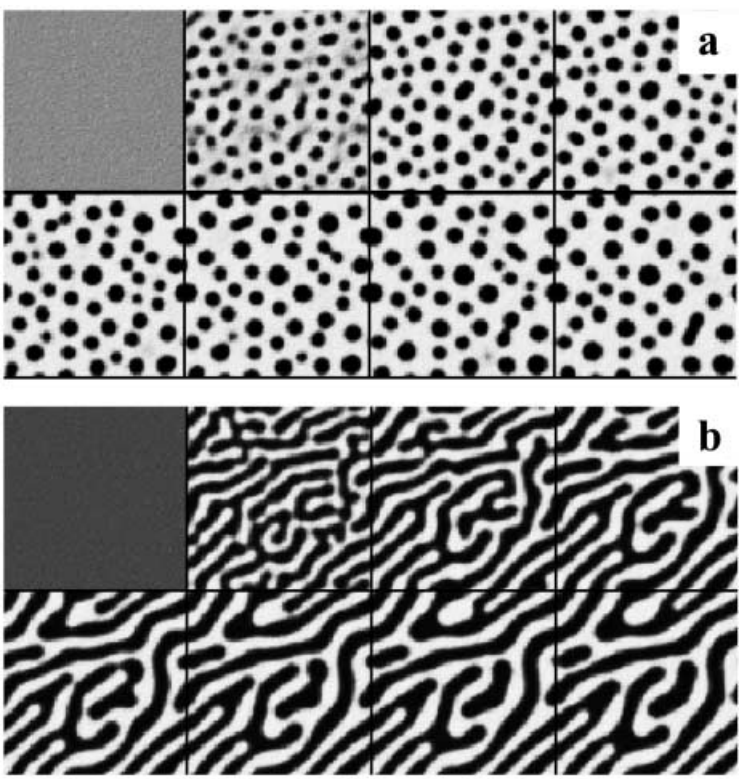

Fig. 2. Two-dimensional phase separation in (6) for $L = 16$, $u_0/k_BT = 6$, $r_0 = 0.2$, $D = 1$, $Z = 1000$, and $< c >= 0.3$ in (**a**) and $< c >= 0.5$ in (**b**). The time intervals between subsequent frames (shown from top left to bottom right) are $\Delta t = 2.5$; gray-scale coding as in Fig. 1

Adsorption and desorption processes modify the properties of the phase transition, because the total number of particles on the surface is not conserved. The kinetics of such a phase transition, combining the features typical for systems

with conserved and non-conserved order parameters, has been considered by us in [20] using equation (1) with $k_r = 0$.

3 Stationary Nonequilibrium Nanostructures

The introduction of a chemical reaction changes the kinetic behaviour of adsorbates undergoing phase transitions. When the particles can arrive by adsorption from the gas phase, compensating for their removal by the reaction, this system can form stationary nonequilibrium nanostructures. The development and the final form of such nanostructures are described by the mesoscopic evolution equation (1).

Neglecting the fluctuations due to internal noises, the stability of the uniform stationary states of equation (1) with respect to small periodic perturbations $\delta c(\mathbf{r}, t) \sim \exp(i\mathbf{kr} + \gamma_k t)$ was analyzed in [22]. This investigation has shown that such states can be unstable with respect to the growth of periodic stationary modes. For sufficiently small interaction radii the wavelength of the first unstable mode can be estimated as

$$\lambda_0 \sim \sqrt{r_0 L_{dif}} \tag{7}$$

where r_0 is the interaction radius and $L_{dif} = \sqrt{D/k_r}$ is the diffusion length of adsorbed reacting particles.

Numerical simulations have shown that this instability leads to the formation of stationary nonequilibrium structures of different morphologies, such as spot arrays and complex labyrinthine phases [22]. The characteristic length scale of such patterns is roughly determined by the wavelength (7) of the first unstable mode. The radius of lateral interactions r_0 is typically a few elementary lattice lengths l_0, i.e. it is much shorter than the diffusion length L_{dif}. Therefore, we have $r_0 \ll \lambda_0 \ll L_{dif}$, and the characteristic length scales of such nonequilibrium reaction-induced stationary structures can be much shorter than the diffusion length and lie in the nanometer range.

Since these structures are so small, fluctuations can play an important role. In numerical simulations they can be taken into account by retaining the noise terms in equation (1). Fig. 3 shows an example of a stochastic simulation of a labyrinthine nanostructure in the presence of internal noises. Four snapshots of the two-dimensional system corresponding to subsequent time moments are displayed here. Measured in elementary lattice lengths l_0, the interaction radius is $r_0 = 6$ and the diffusion length is $L_{dif} = 149$. The total linear size of the system is just $L = 506$ elementary lattice lengths. We see that though the fluctuations are significant, they do not destroy the labyrinthine nanostructure.

The instability responsible for the development of stationary nanostructures is mathematically similar to what is found in reaction-diffusion models supporting Turing patterns. However, there are important physical differences between these two phenomena. The periodic Turing patterns exist in two-component activator-inhibitor systems where the diffusion length $L_{dif,a}$ of the activator is much shorter than the respective diffusion length $L_{dif,i}$ of the inhibitor component. The characteristic wavelength λ of the Turing pattern is estimated as

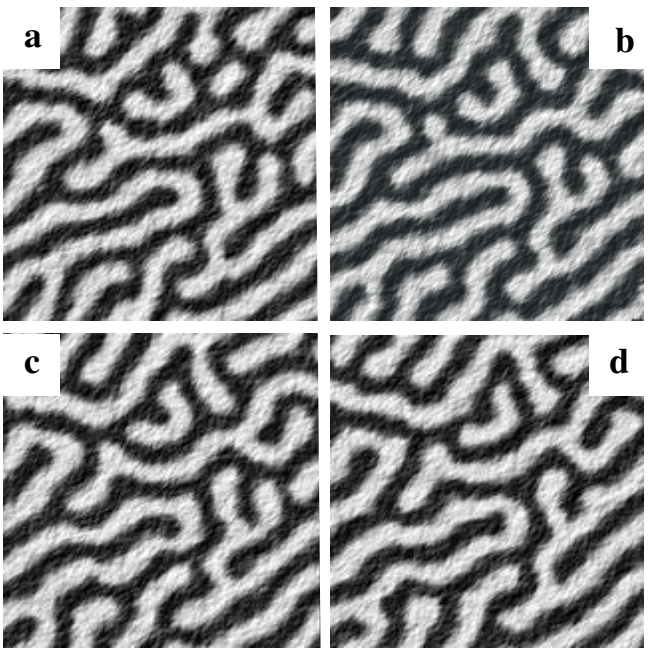

Fig. 3. Fluctuating coverage distributions obtained by numerical integration of the stochastic evolution equation (1) with internal noises (5) for $Z = 2.2 \times 10^3 L_{dif}^{-2}$, system size $L = 3.39\, L_{dif}$, $r_0 = 0.042\, L_{dif}$, $u_0 = 6.4\, k_B T$, $k_r = 0.5\, k_a p_0$, and $k_a p_0 = 0.09\, k_{d,0}$. The time intervals between subsequent frames are $\Delta t = 15/k_{d,0}$; gray-scale coding as in Fig. 1

$\lambda \sim \sqrt{L_{dif,a} L_{dif,i}}$. Hence, it cannot be shorter than the smallest diffusion length $L_{dif,a}$. In contrast to this, stationary reaction-induced nanostructures are found in a single-component system and their characteristic wavelength is shorter than the diffusion length of that component. Moreover, while the formation of a Turing pattern is a purely *kinetic* effect (indeed, only two kinetic diffusion lengths determine their properties), reaction-induced nanostructures result from an interplay between energetic interactions (characterized by the attraction radius r_0) and the kinetic processes (i.e., reaction and diffusion) in the considered system.

On the other hand, the discussed nanostructures are also different from equilibrium microstructures. The latter are found at thermal equilibrium in the systems with short-range attractive and long-range repulsive interactions between particles [7]. Their characteristic wavelength is determined by the combination of two energetic parameters and can often be estimated as $\lambda \sim \sqrt{r_{attr} r_{repuls}}$ where r_{attr} is the (short) radius of attraction and r_{repuls} is the (long) radius of the repulsive interaction between two particles.

Thus, it can be concluded that stationary reaction-induced nanostructures represent a physically different kind of patterns, occupying an intermediate po-

sition between purely kinetic Turing structures and equilibrium microstructures of a purely energetic origin.

4 Traveling Nanostructures

It is generally known that, in contrast to equilibrium systems where only stationary structures can be observed, nonequilibrium physical systems can also show regular or chaotic self-oscillations and may support the active propagation of waves. Many different traveling wave patterns are known in the theory of reaction-diffusion systems. However, the smallest characteristic wavelength of such traveling patterns is again limited by the diffusion length and cannot be very short. Such a limitation is absent for systems combining reactions and diffusion with potential energetic interactions between particles. Traveling nanostructures, that may contain as few as only hundreds of reacting atoms, are possible under these conditions.

As an example [24], we consider a surface chemical reaction with two reacting species U and V. Both species arrive through adsorption from the gas phase. However, only the adsorbed molecules of species U are mobile and can diffuse over the surface. When they meet immobile adsorbed molecules V, both molecules react and give rise to a product that immediately leaves the surface ($U + V \to 0$). We assume that attractive lateral interactions between molecules U and also between molecules U and V are present. The mesoscopic kinetic equations for fluctuating local coverages u and v of both species are derived from the master equation for the full probability distribution of such a reactive lattice gas (see [25]). They have the form

$$\frac{\partial u}{\partial t} = k_{a,u} p_u (1-u) - k_d u - k_r uv + \frac{\partial}{\partial \mathbf{r}} \left(\frac{D}{k_B T} (1-u) u \frac{\partial W}{\partial \mathbf{r}} \right) + D \frac{\partial^2 u}{\partial \mathbf{r}^2} + \xi_u(\mathbf{r}, t)$$

$$\frac{\partial v}{\partial t} = k_{a,v} p_v (1-v) - k_r uv + \xi_v(\mathbf{r}, t) \tag{8}$$

The local surface potential $W(\mathbf{r})$ is determined by attractive interactions between particles,

$$W(\mathbf{r}) = -\int w_{uu}(\mathbf{r} - \mathbf{r}') u(\mathbf{r}') d\mathbf{r}' - \int w_{uv}(\mathbf{r} - \mathbf{r}') v(\mathbf{r}') d\mathbf{r}' \tag{9}$$

For simplicity, we assume that both pair attraction potentials $w_{uu}(r)$ and $w_{uv}(r)$ have form (3) with the same interaction radius r_u and different intensities w^0_{uu} and w^0_{uv}. The desorption rate k_d of molecules U depends on the local potential as $k_d = k_{d,0} \exp[W(\mathbf{r})/k_B T]$; the immobile molecules V do not desorb. The reaction rate k_r does not depend on the local potential since, as we assume, the chemical energy released in an individual reaction event is sufficiently high to overcome any local differences in the adsorption potential. The internal noises in these two equations are

$$\xi_u(\mathbf{r}, t) = \frac{1}{Z^{1/2}} \sqrt{k_{a,u} p_u (1-u)} f_{a,u}(\mathbf{r}, t) + \frac{1}{Z^{1/2}} \sqrt{k_d} f_d(\mathbf{r}, t)$$

$$+\frac{1}{Z^{1/2}}\sqrt{k_r uv}f_r(\mathbf{r},t)+\frac{1}{Z^{1/2}}\frac{\partial}{\partial \mathbf{r}}\left(\sqrt{2Du(1-u)}\mathbf{f}(\mathbf{r},t)\right) \qquad (10)$$

$$\xi_v(\mathbf{r},t)=\frac{1}{Z^{1/2}}\sqrt{k_{a,v}p_v(1-v)}f_{a,u}(\mathbf{r},t)+\frac{1}{Z^{1/2}}\sqrt{k_r uv}f_r(\mathbf{r},t) \qquad (11)$$

where $f_{a,u}, f_d, f_r, f_{a,v}$ and $\mathbf{f}$ are independent white random forces of unit intensity. Note that the reaction noise in both equations is exactly the same because each reaction event involves molecules U and V.

Analyzing the stability of uniform states with respect to small periodic perturbations, we have found [24] that for sufficiently small interaction radii the first unstable mode in this system has a finite wavelength $\lambda_0 \sim \sqrt{r_0 L_{dif}}$. However, in contrast to the situation typical for the previously discussed model, the growing mode is now oscillatory (i.e. $\mathrm{Im}\gamma_k \neq 0$). Therefore, the instability represents a Hopf bifurcation with finite wave number (also called a "wave bifurcation"). Such bifurcations are known in reaction-diffusion systems where they give rise to patterns of traveling or standing waves (see, e.g. [26]).

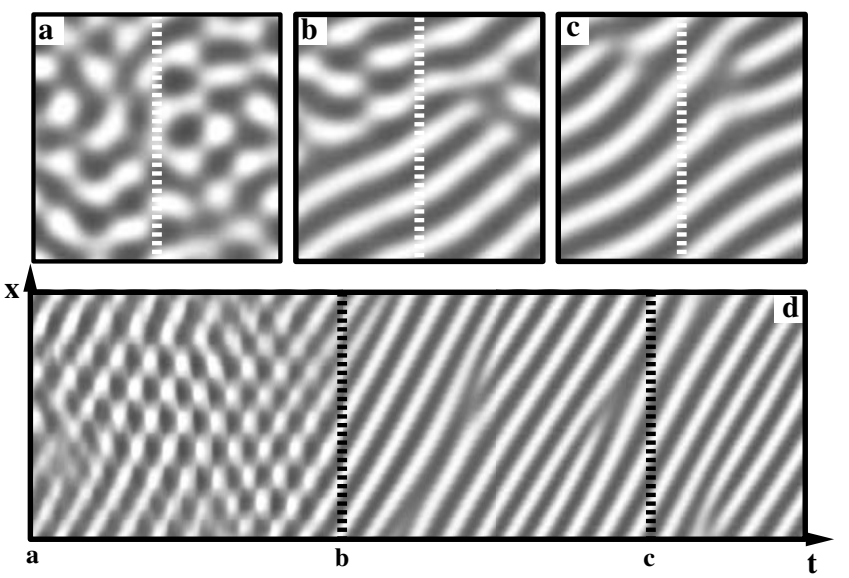

Fig. 4. Formation of coherent traveling wave patterns in a two-dimensional simulation of (8) in the absence of internal fluctuations for $w^0_{uu} = 5\,k_BT$, $w^0_{uv} = 3\,k_BT$, $k_{a,u}p_u = k_r = 0.5\,k_{d,0}$, $k_{a,v}p_v = 1.5\,k_{d,0}$, and $r_0 = 0.07\,L_{dif}$. The snapshots (**a**, **b**, **c**) correspond to time moments (**a**) $t = 18/k_{d,0}$, (**b**) $t = 45/k_{d,0}$, and (**c**) $t = 72/k_{d,0}$. The temporal evolution is displayed in (**d**) during the interval $18/k_{d,0} < t < 85/k_{d,0}$ in the one-dimensional cross-section indicated by the vertical dashed lines in the snapshots. The total size of the system is $L = 4.2\,L_{dif}$. The local surface coverage of the species U is shown in gray scale, with darker areas corresponding to higher coverages

Numerical simulations of equations (8) in absence of noises indeed show that, depending on the parameters, either standing waves or traveling waves are observed in the system. As an example, Fig. 4 displays the results of a typical two-dimensional simulation, starting with small random perturbations added to the unstable uniform phase. The three snapshots (a,b,c) show the coverage distributions of the species U at subsequent time moments. To characterize the dynamics of the pattern, Fig. 4d presents its temporal evolution along the central cross section shown by the dashed lines in these three snapshots. It can be seen that, at the early stage of evolution, irregular patterns of distorted standing waves are formed. Later, however, the standing waves evolve into planar wave trains traveling at a constant velocity. These waves tend to form a periodic pattern containing pointlike defects. For other parameter choices, stable standing waves were found in numerical simulations [25].

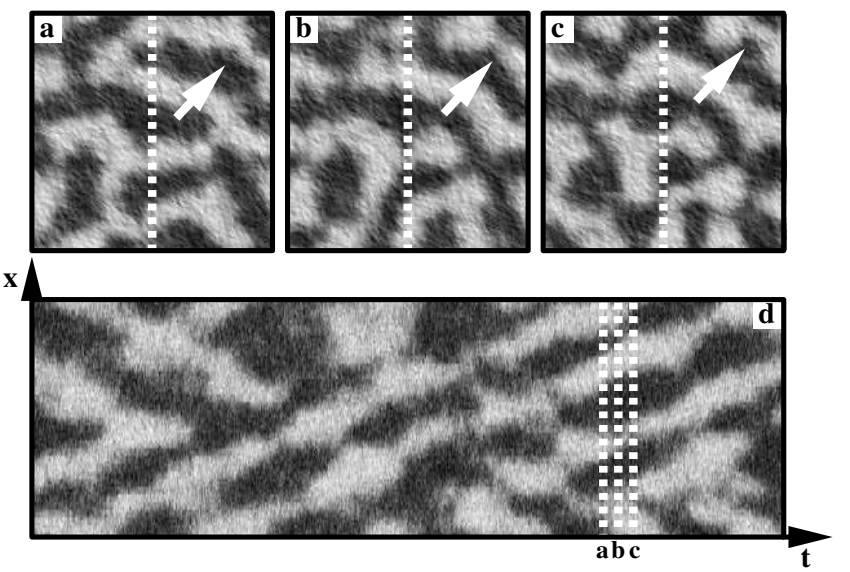

Fig. 5. Fluctuating traveling wave fragments as obtained by numerical integration of (8) with internal fluctuations given by (10) and (11); $Z = 1.07 \times 10^5\, L_{dif}^{-2}$, $k_{a,u} p_u = 0.08\, k_{d,0}$, $L = 1.7\, L_{dif}$, $r_0 = 0.028\, L_{dif}$; the other parameters are the same as in Fig. 4. The two-dimensional snapshots (**a**, **b**, **c**) are separated by equal intervals $\Delta t = 0.89/k_{d,0}$. The temporal evolution in the one-dimensional cross-section (**d**) is shown during time $t = 44.6/k_{d,0}$; gray-scale coding as in Fig. 4

Since the structures are very small and may include only relatively small number of adsorbed molecules, internal fluctuations should be taken into account. Fig. 5 shows the fluctuating coverage distributions obtained by numerical integration of equations (8) with the noise terms (10) and (11). Here, the total size of the system is only $L = 555$ elementary lattice lengths l_0, the interaction

radius is $r_0 = 9\,l_0$ and the diffusion length is $L_{dif} = 327\,l_0$. Clearly, at these small scales the internal noises of the diffusion, adsorption, desorption and reaction processes exhibit strong influence on the patterns. The individual traveling stripes are now broken into many short fragments that form irregular spatial patterns seen in the snapshots (a,b,c). Nonetheless, examining the time evolution in the central cross section (Fig.5d), we recognize that these fragments do not just fluctuate. They travel across the surface while undergoing irregular variations of their shapes. Merging of traveling fragments, as well as splitting, is observed in the process. Remarkably, the magnitude of the propagation velocity of different fragments does not significantly differ.

Though traveling nanoscale structures were found by us only in a particular model of a surface chemical reaction, the mechanisms responsible for their formation are general and similar structures can therefore be expected in other nonequilibrium soft matter, such as reactive polymer blends or biological membranes. These extremely small traveling structures would perfectly fit into the characteristic dimensions of a single biological cell.

It should be noted that the propagation of these nanostructures involves mass transport through the medium. Essentially, they represent traveling atomic clusters. Hence, they can be used for the transport and the delivery of particles.

5 Self-Organized Chemical Nanoreactors

As noted in Section 2, adsorbates can also induce structural phase transitions in the top crystalline layer of the metal substrate. Since the structural state of the surface determines the local adsorption energy, such a phase transition influences the diffusive motion of adsorbed particles and the course of a surface chemical reaction. This can lead to the formation of localized structures representing self-organized chemical nanoreactors [27].

Interactions between the adsorbed particles are not taken into account below. We assume that the structural state of the surface can be characterized by a continuous order parameter η and that the free energy of the system is given by the Ginzburg-Landau functional

$$F[\eta, c] = \int \left(\frac{1}{2}\kappa^2 \left(\nabla\eta\right)^2 + \varphi(\eta, c) + h(c) \right) d\mathbf{r} \tag{12}$$

where the coefficient κ specifies the strength of the coupling between phase transitions in neighbouring elements of the surface,

$$h(c) = k_B T n_0 \left[(1-c)\ln(1-c) + c\ln c\right] \tag{13}$$

is free energy density of the ideal lattice gas of adsorbed particles (n_0 is the number of adsorption sites per unit area) and $\varphi(\eta, c)$ is the density of the free energy associated with the structural phase transition. We choose it as a polynomial of the fourth order, i.e.

$$\varphi(\eta, c) = \chi \left[\eta^4 - \frac{8}{3}\eta^3 + 2\eta^2 + c\left(\frac{4}{3}\eta^3 - 2\eta^2\right) \right] \tag{14}$$

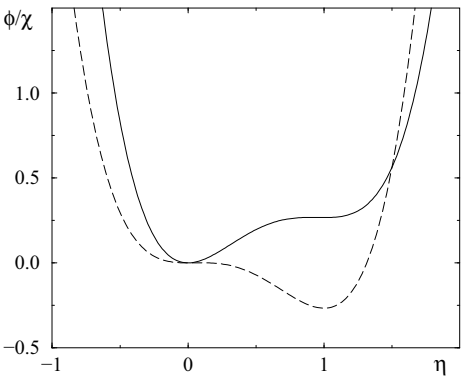

Fig. 6. Normalized free energy density $\phi(\eta, c)/\chi$ as given by (14) for a homogeneous distribution of the order parameter η and two different values of the uniform adsorbate coverage $c = 0.1$ (*solid line*) and $c = 0.9$ (*dashed line*)

Fig. 6 shows the free energy density $\varphi(\eta, c)$ for two different values of the coverage c. When $c < 0.5$, the deepest energy minimum (and thus the absolutely stable state) corresponds to $\eta = 0$. If the coverage c is increased and exceeds $c = 0.5$, the deepest energy minimum becomes located at $\eta = 1$. Hence, by changing the adsorbate coverage the structural state of the surface can be controlled.

If the order parameter of the considered phase transition is not conserved, its relaxation kinetics should be described by

$$\frac{\partial \eta}{\partial t} = -\Gamma \frac{\delta F[\eta, c]}{\delta \eta(\mathbf{r}, t)} \tag{15}$$

or, explicitly,

$$\tau \frac{\partial \eta}{\partial t} = \eta(1 - \eta)(\eta + c - 1) + l^2 \frac{\partial^2 \eta}{\partial \mathbf{r}^2} \tag{16}$$

where $\tau = (4\chi\Gamma)^{-1}$ is the characteristic relaxation time of the phase transition and $l = \kappa/\sqrt{4\chi}$ is the associated microscopic length scale (determining the width of an interface between two phases).

The motion of adsorbed particles over the surface and their desorption are influenced by the local chemical potential of the surface,

$$\mu(\mathbf{r}) = \frac{1}{n_0} \frac{\partial}{\partial c} \left[\varphi(\eta, c) + h(c)\right] = U(\eta) + k_B T \ln\left(\frac{c}{1 - c}\right) \tag{17}$$

where U is the local potential associated with the phase transition,

$$U(\eta) = -\frac{2\chi}{n_0}(\eta^2 - \frac{2}{3}\eta^3) \tag{18}$$

We see that in structural state $\eta = 0$ of the surface this potential is $U = 0$, whereas in the other structural state $\eta = 1$ it takes the value $U = -2\chi/3n_0$.

Thus, the surface regions occupied by the phase with $\eta = 1$ will represent potential wells for diffusing adsorbed particles and will trap them if these wells are sufficiently deep.

The evolution of the adsorbate coverage, taking into account adsorption, desorption, reaction and diffusion on the surface, is governed by the general kinetic equation

$$\frac{\partial c}{\partial t} = k_a p_0 (1-c)\left[1 - \exp\left(\frac{\mu(\mathbf{r}) - \mu_0}{k_B T}\right)\right] + \frac{\partial}{\partial \mathbf{r}}\left[\frac{D}{k_B T} c(1-c)\frac{\partial \mu}{\partial \mathbf{r}}\right] - k_r c \quad (19)$$

where $\mu_0 = k_B T \ln(k_a p_0 / k_{d,0})$ is the equilibrium chemical potential, determined by the balance of the adsorption and the desorption processes (p_0 is the partial pressure in the gas phase above the surface, k_a is the sticking coefficient and $k_{d,0}$ is the desorption rate constant for the surface structural state with $\eta = 0$). We assume that the reaction is nonequilibrium (e.g. photoinduced) and therefore its rate is not influenced by the local surface potential.

Substituting the chemical potential given by equation (17) and performing some transformations, we finally obtain the following kinetic equation

$$\begin{aligned}\frac{\partial c}{\partial t} &= k_a p_0 (1-c) - k_{d,0} \exp\left(\frac{U(\eta)}{k_B T}\right) c - k_r c \\ &+ \frac{\partial}{\partial \mathbf{r}}\left[\frac{D}{k_B T} c(1-c)\frac{\partial U(\eta)}{\partial \mathbf{r}}\right] + D\frac{\partial^2 c}{\partial \mathbf{r}^2}\end{aligned} \quad (20)$$

Note that the derived kinetic equation contains not only the diffusion term, but also a term that describes a viscous adsorbate flow induced by the potential gradient.

Equations (16) and (20) form the mathematical model of the considered system. It should be noted that the diffusion length $L_{d,0} = \sqrt{D/k_{d,0}}$ of adsorbed particles is much larger than the microscopic characteristic length l of the phase transition. Therefore, an investigation based on the singular perturbation theory is possible. The analysis in [27] shows that this system has stationary localized solutions which represent islands of the phase $\eta = 1$ which are surrounded by the phase with $\eta = 0$. Since the attraction potential U inside the islands is much stronger, they trap diffusing particles and their concentration inside the islands is high ($c \simeq 1$) This means that the chemical reaction will mainly go on inside such small surface regions, so that the reactive islands would represent self-organized chemical microreactors.

Figure 7 shows the computed dependence of the radius R of stationary islands on the reaction rate constant k_r. We see that localized structures become possible only when a certain critical reaction rate is exceeded. The dependence has two branches, the lower of which (dashed line) is always unstable. The radius of the island decreases with the reaction rate and may become much smaller than the diffusion length for adsorbed particles.

The stability of the islands can first be discussed using simple arguments. Suppose that we have an island of radius R containing N adsorbed particles.

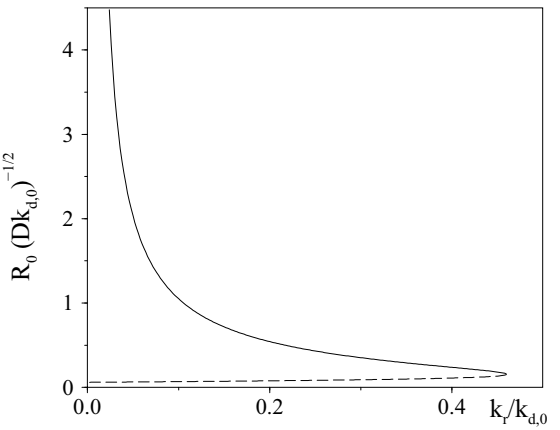

Fig. 7. Dependence of the radius R of stationary circular spots on the dimensionless reaction rate constant $k_r/k_{d,0}$. The other parameters are $k_a p_0 = 0.05\, k_{d,0}$, $\chi = 15\, n_0\, k_B T$, $\tau = (15\, k_{d,0})^{-1}$, and $l = 0.032\, \sqrt{D/k_{d,0}}$

The number N will increase with time because of the diffusion flux of adsorbed particles arriving into the island. This flux is proportional to the length of the island boundary and hence to the radius R. On the other hand, the number of particles inside the island will decrease with time because of the reaction (desorption from the island can be neglected). The respective flux is proportional to the island area, i.e. to R^2. Hence, we have

$$\frac{dN}{dt} = \alpha R - \beta R^2 \tag{21}$$

with positive coefficients α and β. Assuming that inside the island the maximal coverage $c = 1$ is reached, we have $N = \pi R^2 n_0$. Therefore, an approximate evolution equation for the island radius is obtained,

$$2\pi n_0 R \frac{dR}{dt} = \alpha R - \beta R^2 \tag{22}$$

It can easily be checked that it has a stable stationary solution with $R = \alpha/\beta$.

The above arguments indicate that the islands will be stable with respect to radially symmetric perturbations. However, they do not exclude instabilities involving shape deformations of the islands. Our numerical simulations reveal that small reactive islands with the radius close to the microscopic length l are always stable [27]. Large islands are unstable and give rise to growing snowflake structures. An example of such a developing structure is shown in Fig. 8. The further development of such instabilities may lead to the replication of the localized structures or to the formation of an extended labyrinthine pattern.

The origin of the instability is related to the fact that the adsorbate coverage is depleted by diffusion in the lagunas formed by the concave part of the island boundary. Therefore the interface velocity in such regions is decreased as compared with that in the protruding parts where the adsorbate coverage is higher.

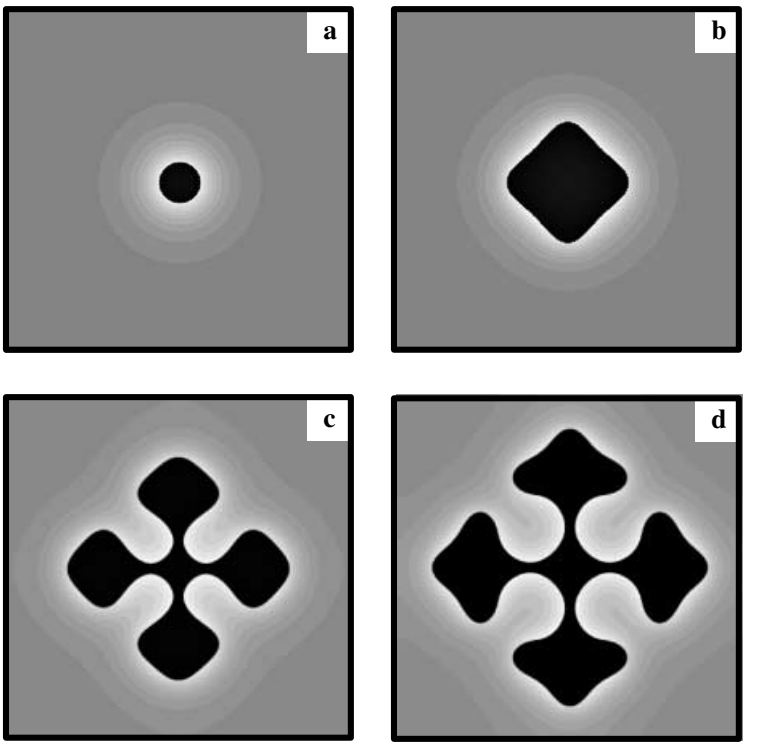

Fig. 8. Instability of a circular spot in a two-dimensional numerical simulation of (16) and (20). Frames (**a**) - (**d**) correspond to the time moments $t = 0$, $65/k_{d,0}$, $237/k_{d,0}$, and $363/k_{d,0}$ respectively; $k_r = 0.05\,k_{d,0}$, and the other parameters were chosen as in Fig. 7. The local coverage c is shown in grey scale, increasing from white to black

A similar instability is encountered in the dendritic growth of crystals [28] and in biological chemotaxis problems [29].

In contrast to prefabricated microreactors, self-organized nanoreactors exist only if a chemical nonequilibrium reaction is present. Their sizes may be easily controlled by changing the reaction parameters. Such nanoreactors are dynamical objects and modifications of the considered simple model can probably yield localized nanostructures that would breath or travel over the surface.

References

1. R. Kapral, K. Showalter (eds.): *Chemical Waves and Patterns* (Kluwer, Dordrecht 1995)
2. A. S. Mikhailov: *Foundations of Synergetics I* (Springer, Berlin 1990, 2nd revised edition 1994)
3. B. Hess, A.S. Mikhailov: Science **264**, 223 (1994)
4. A. S. Mikhailov, G. Ertl: Science **272**, 1596 (1996)

5. J. Lechleiter, S. Girard, E. Peralta, D. Clapman: Science **252**, 123 (1991)
6. H. R. Petty, R. G. Worth, A.L. Kindzelskii: Phys. Rev. Lett. **84**, 2754 (2000)
7. M. Seul, D. Andelman: Science **267**, 476 (1995)
8. S. C. Glotzer, E. A. Di Marzio, M. Muthukumar: Phys. Rev. Lett. **74**, 2034 (1995)
9. M. Motoyama, T. Ohta: J. Phys. Soc. Jpn., **66**, 2715 (1997)
10. Q. Tran-Cong, A. Harada: Phys. Rev. Lett. **76**, 1162 (1996)
11. Y. Tabe, H. Yokoyama: Langmuir **11**, 4609 (1995)
12. Ch. D. Zangmeister, J. E. Pemberton: J. Phys. Chem. B **102**, 8950 (1998)
13. J. Trost, T. Zambelli, J. Wintterlin, G. Ertl: Phys. Rev. B **54**, 17850 (1996)
14. J. Wintterlin, S. Völkening, T.V.W. Janssens, T. Zambelli, G. Ertl: Science **278**, 1931 (1997)
15. T.V.W. Janssens, S. Völkening, T. Zambelli, J. Wintterlin, G. Ertl: J. Phys. Chem. **B 102**, 6521 (1998)
16. S. Völkening, K. Bedürftig, K. Jacobi, J. Wintterlin, G. Ertl: Phys. Rev. Lett. **83**, 2672 (1999)
17. S. Renisch, R. Schuster, J. Wintterlin, G. Ertl: Phys. Rev. Lett. **82**, 3839 (1999)
18. E. Bauer, in: 'Structure and Dynamics of Surfaces II, Topics in Current Physics Vol. 43', ed. by W. Schommers, P. v. Blanckenhagen (Springer, Berlin, Heidelberg 1987).
19. A. S. Mikhailov, G. Ertl: Chem. Phys. Lett. **238**, 104 (1995); **267**, 400 (1997)
20. M. Hildebrand, A.S. Mikhailov: J. Phys. Chem. **100** , 19089 (1996)
21. J. W. Cahn: Acta Metall. **9**, 795 (1961)
22. M. Hildebrand, A.S. Mikhailov, G. Ertl: Phys. Rev. E **58**, 5483 (1998)
23. F. Ulmar: Diploma Thesis, Department of Physics, Technical University, Berlin (1998)
24. M. Hildebrand, A. S. Mikhailov, G. Ertl: Phys. Rev. Lett. **81**, 2602 (1998)
25. M. Hildebrand, A.S. Mikhailov: J. Stat. Phys., in press
26. D. Walgraef: *Spatio-Temporal Pattern Formation* (Springer, Berlin 1997)
27. M. Hildebrand, M. Kuperman, H. Wio, A.S. Mikhailov, G. Ertl: Phys. Rev. Lett. **83**, 1475 (1999)
28. J.S. Langer: Rev. Mod. Phys. **52**, 1 (1980)
29. M. Mimura, T. Tsujikawa: Physica A **230**, 499 (1996)

Pattern Formation in Electric Discharges

Ute Ebert[1] and Manuel Arrayás[1,2]

[1] CWI, P.O.Box 94079, 1090 GB Amsterdam, The Netherlands
[2] Instituut-Lorentz, P.O.Box 9506, 2300 RA Leiden, The Netherlands

Abstract. Electric discharges with low current densities can be described by a minimal model of reaction, drift and diffusion of two charged species coupled nonlinearly to an externally imposed electric field. The intrinsic ionization–field–coupling of the model leads to pattern formation both in stationary or periodically driven barrier discharges and in transient streamer discharges. In particular, we discuss negative streamer discharge fingers extending into a non–ionized area. They belong to the class of Laplacian growth phenomena. For quantitative predictions, a systematic derivation of the corrections to the velocity of a moving boundary due to curvature or relaxation effects is still missing. This is because negative streamer fronts are so–called pulled fronts, for which standard approximation schemes fail. We present new approaches and results.

1 Modelling Electric Discharges

1.1 Phenomenology, Spatio-temporal Patterns

Electric discharges are commonly known from natural phenomena like sparks whose lengths can vary from cm's to km's, or St. Elmo's fire (a corona discharge at the tops of ship masts). Neon tubes are the best known example from a whole list of technical applications that are under continued use and investigation. For an overview of phenomena in gas discharges, we refer to [1]. Discharges can occur not only in gases, but also in fluids or solids – in just any matter that can turn from a state of low or vanishing conductivity to a state of high conductivity, when a sufficiently strong field is applied.

The name "discharge" is a historical one like many other ones in this field, since it refers to the discharge of the voltage stored in a capacitor, when the matter between the capacitor plates becomes conducting. However, a sufficiently strong voltage source also might sustain the voltage difference between the outer electrodes despite the "discharge" in the bridging matter. Electric discharges can create a *low temperature plasma* locally. In contrast to high temperature plasmas as in fusion reactors or stars that exist due to heating and confinement, low temperature plasmas exist under nonequilibrium conditions due to external excitations like electromagnetic fields. Hence they generically are inhomogeneous in space and time and form spatio-temporal patterns. The homogeneous glow of the neon tube is an exception and a result of the art of the engineer. The onset of spatio-temporal instabilities limits the range of technical applicability in quite a number of cases.

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 270–282, 2001.
© Springer-Verlag Berlin Heidelberg 2001

Below, we will focus on *streamer and barrier discharges* in simple matter like pure nitrogen under normal conditions. Streamers are the initial breakdown mode of insulating matter on length scales from cm's to m's after a voltage shock has been applied. They are extending finger–like patterns of ionized matter, and also the precursors of a later short–circuit, if the conducting channel eventually contacts both electrodes. Initially during the streamer phase, the degree of ionization stays relatively low.

Barrier discharges are operated in a stationary or periodically changing electric field. They consist of a layered structure of a discharge with at least one resistive layer in dc fields or at least one dielectric layer in ac fields. High currents and degrees of ionization are then prevented by the dielectric or resistive layers, just as in the case of a large external load resistance in the electric circuit.

1.2 The Minimal Model of Low Current Discharges and the Ionization–Field–Coupling

Pattern formation occurs in nonlinear spatially extended systems under nonequilibrium conditions. The nonequilibrium here is due to an externally applied electric field. The nonlinear mechanism to be explored is an *ionization–field–coupling*: a sufficiently high electric field leads to a multiplication of charge carriers by a local impact ionization reaction. The generated charges drift in the electric field. If their density is sufficiently high, they will modify the field and hence change the local reaction rates and drift velocities. Accordingly, the evolution of the local degree of ionization is determined by a fully dynamic process. The situation is captured by the following model:

In general, the continuity equations for particles of species i are

$$\partial_t \rho_i + \nabla \cdot \mathbf{j}_i = f_i\left(\{\rho_j\}, \mathbf{E}\right) \ , \quad \mathbf{j}_i = q_i \mu_i \rho_i \mathbf{E} - D_i \nabla \rho_i \ , \tag{1}$$

where ρ_i is the particle density, $q_i = \pm$ the sign of the charge, μ_i the mobility and D_i the diffusion constant of species i. If the degree of ionization is sufficiently small, dissipative heating can be neglected and the neutral background stays unchanged. The current $\mathbf{j}_i$ then can be approximated as in (1) by particle diffusion and Ohmic friction. This is true in gases, liquids or solids. In a semiconductor, the current $\mathbf{j}_i$ can saturate for large $\mathbf{E}$. f_i is the source term describing the reactions of the particles in the local field. If the currents stay sufficiently small, also magnetic and relativistic effects can be neglected, and we only need the Poisson equation of electrostatics

$$\frac{q}{\varepsilon_0} = \sum_i \frac{q_i e \ \rho_i}{\varepsilon_0} = \nabla \cdot \mathbf{E} \quad , \quad \mathbf{E} = -\nabla \Phi \ , \tag{2}$$

where e is the electron unit charge and ϵ_0 the dielectric constant. The dominant reaction in a strong electric field is impact ionization: a light charged particle, typically an electron e^-, can react with a neutral atom or molecule A as: $\mathrm{e}^- + \mathrm{A} \longrightarrow 2\mathrm{e}^- + \mathrm{A}^+$, if local electric field times mean free path of the electron are of the order of the ionization energy.

In the simplest case (for a more extended discussion, see [2,3]), there are only two charged species in a gas, namely electrons (denoted by $i = e$) and positive ions ($i = +$), and the reaction terms f_i consist of impact ionization and recombination only:

$$f_e(\rho_e, \rho_+, \mathbf{E}) = f_+ = |\mu_e \rho_e \mathbf{E}|\ \alpha(|\mathbf{E}|) - \beta \rho_e \rho_+ \ . \tag{3}$$

Due to charge conservation, the two reaction rates f_e and f_+ are identical. The effective cross-section $\alpha(|\mathbf{E}|)$ has approximately a threshold dependence on the electric field. The traditional Townsend approximation [1] reads, e.g.

$$\alpha(|\mathbf{E}|)_{\text{Townsend}} = \alpha_0\ e^{-E_0/|\mathbf{E}|}\ . \tag{4}$$

The homogeneous equilibrium conductivity due to the electrons for a given field $\mathbf{E}$ then is $e\mu_e^2|\mathbf{E}|\alpha(|\mathbf{E}|)/\beta$, and hence the equilibrium current $j(E)$ is proportional to $|\mathbf{E}|^2\alpha(|\mathbf{E}|)$. In this model, j is a monotonically increasing function of E, so there is no bistability of the current–voltage–characteristics on the local level of description. Therefore the well–known instability mechanism [4] due to S– or N–shaped current–voltage–characteristics does not apply to simple nitrogen.

If furthermore bulk recombination can be neglected because of a separation of time scales in the streamer discharge or because of the dominance of boundary recombination processes in the barrier discharge, and if the mobility of the ions is neglected because $\mu_+ \ll \mu_e$ and $D_+ \ll D_e$, we arrive at the *minimal model of low current discharges*. In dimensionless form, it reads [5,2]:

$$\partial_t\ \sigma - \nabla \cdot (\sigma\ \mathbf{E} + D\ \nabla\sigma) = \sigma\ f(|\mathbf{E}|)\ , \quad \text{with, e.g. } f(E) = |E|\ e^{-1/|E|}\ , \tag{5}$$

$$\partial_t\ \rho \qquad\qquad = \sigma\ f(|\mathbf{E}|)\ , \tag{6}$$

$$\nabla \cdot \mathbf{E} \qquad = \rho - \sigma\ , \qquad \mathbf{E} = -\nabla\Phi\ . \tag{7}$$

σ and ρ are the electron and ion particle densities, and $\mathbf{E}$ is the electric field. The dimensionless diffusion constant $D = D_e\alpha_0/(\mu_e E_0)$ is the only intrinsic parameter. For nitrogen under normal conditions the units in (5)–(7) are [2,6]: 2 μm for length, 3 ps for time t, and hence c/400 (c = speed of light) for velocity v, furthermore 200 kV/cm for field $|\mathbf{E}|$, and $5 \cdot 10^{14}/\text{cm}^3 \approx 2 \cdot 10^{-5}\rho_0$ for the particle densities σ and ρ, where ρ_0 is the particle density of the neutral N_2 under normal conditions. The diffusion constant is $D = 0.1$.

The minimal model applies to non-attaching gases like N_2 or He. In attaching gases like O_2, negative ions need to be included that can be formed by the attachment reaction: $e^- + O_2 \longrightarrow O_2^-$. In doped semiconductors, there are generally also three types of charged species: (*i*) mobile electrons, (*ii*) mobile electron holes and (*iii*) either immobile donors or immobile acceptors. In the reaction term (3), impact ionization by electron holes also needs to be taken into account. In direct band semiconductors such as, e.g., GaAs or GaN, or in N_2/O_2 mixtures, the possible importance of photoionization processes requires additional attention.

1.3 Pattern Formation in Barrier Discharges

Before embarking into a more extensive discussion of streamer discharges, we briefly discuss barrier discharges with stationary (dc) or periodical (ac) drive as another manifestation of the ionization–field–coupling.

In particular, planar layered systems with a sequence like (*i*) transparent electrode, (*ii*) gas discharge, (*iii*) resistive or dielectric layer (for dc or ac voltage, resp.), and (*iv*) the second electrode are under continued careful experimental investigation [7,8]. The discharge and the linear layer are orders of mm thick, and the whole structure has a wide lateral aspect ratio. Looked at from above through the transparent electrode, these systems exhibit an extremely rich variety of effectively two–dimensional patterns in the plane orthogonal to the layers. What is visible, is the light emitted from the ionized areas which is the sum of emissions over the whole height of the discharge layer.

The patterns resemble other effectively two–dimensional systems like Rayleigh–Bénard–convection or electroconvection in liquid crystals. The existence of a supercritical and a weakly subcritical bifurcation with associated amplitude patterns has been established [7], as well as a whole zoo of particle–like excitations with most striking interactions, replications and annihilations [7,8].

However, a derivation of the patterns from gas discharge equations of type (5)–(7) has long been missing. An early attempt to predict the dc patterns [9] is based on the ad–hoc assumption of a locally bistable current–voltage–characteristics, which is common in semiconductor systems [4], but not applicable to discharges in nitrogen, as was discussed above. On the other hand, bistability of the discharge on a global level can result from the ionization–field–coupling. For appropriate boundary conditions, this instability can lead to the formation of a glow discharge with its characteristic space charges [1]. Pattern formation in the dc barrier discharge according to this mechanism is presently under investigation [10]. The starting point is a model like (5)–(7), where the mobility of the ions has to be included, but diffusion can be neglected.

In the case of ac barrier discharges [8], a field–ionization driven mechanism of pattern formation first was suggested in [11]. Indeed, recent numerical investigations [8] of a model of type (5)–(7) could verify this picture.

2 Pattern Formation in Streamer Discharges

2.1 Basic Observations

From here on we focus on the transient process of dielectric breakdown by the streamer mechanism. Streamers occur, when a sufficiently high electric field is suddenly applied to an ionizable medium with low initial conductivity in a sufficiently long discharge gap. The effect was first studied by Raether [12], who experimentally observed a qualitative change of form and velocity of an impact ionization avalanche after a certain time of evolution. This avalanche–to–streamer–transition, he attributed to the space charges becoming important above a critical degree of ionization.

In terms of the minimal model, the *avalanche* is the regime of low particle densities. It amounts to neglecting the space charges $\rho - \sigma$ in (7). The field $\mathbf{E} = -\nabla\Phi$ then is completely determined by the boundary conditions and the Laplace equation $\nabla^2\Phi = 0$. Eqs. (5) and (6) then are linear equations for the densities σ and ρ, which have to be solved in the presence of a given external electric field. In the *streamer* regime, the contributions of the space charges $\rho - \sigma = \nabla \cdot \mathbf{E}$ (7) to the electric field become comparable to the external electric field. The ionization–field–coupling that then couples the three scalar fields σ, ρ and Φ, makes the full model (5)–(7) deeply nonlinear.

2.2 Streamers of Different Flavors

In [13,6], the evolution of streamer discharges from initial ionization seeds was simulated for N_2 under normal conditions between planar electrodes with a distance of 5 mm and a voltage of about 25 kV. The three dimensional problem was reduced to an effectively two dimensional one by the assumption of radial symmetry. Fig. 1 shows two time steps with $\Delta t = 0.75$ ns of the evolution of a negative streamer from Vitello *et al.* [6] – a negative or anode directed streamer propagates in the drift direction of the electrons, a positive or cathode directed streamer antiparallel to the electron drift. Each line in Fig. 1 denotes the increase of electron density by a factor of 10. Densities from 10^3 to 10^{14} free electrons/cm^3 can be seen. Thus the ionization increases by more than 10^{10} within a layer of a few μm's thick.

In dimensionless units, these simulations amount to numerical solutions of the model (5)–(7) with the dimensionless external field $E \approx 1/4$ and electrode distance $L \approx 250$. The dimensionless time difference is $\Delta t \approx 250$.

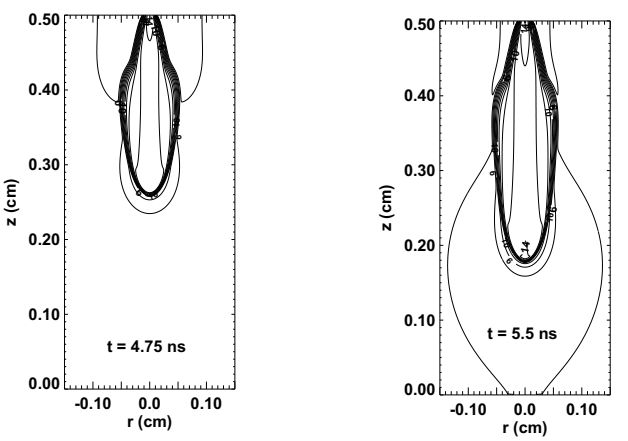

Fig. 1. Figures from [6].: Two time steps of the evolution of a negative streamer into an area with very low background ionization [Courtesy of P.A. Vitello]

The interior of the streamer finger is an area of high ionization and low field, so the ionization reaction $\sigma f(|\mathbf{E}|)$ is suppressed, because $f(|\mathbf{E}|) \approx 0$. There is a well pronounced propagating front regime surrounding the finger, where reaction, drift, diffusion and electric screening take place. In the outer area, the reaction rate $\sigma f(|\mathbf{E}|)$ again is small, now because σ is small. The two simulations [13,6] differ by the properties of the outer area, which are determined by different initial conditions: In the earlier simulations of Dhali and Williams [13], the initial background ionization is 10^8 – $10^{11}/\mathrm{cm}^3$ ($= 2 \cdot 10^{-7}$ – 10^{-4} in dimensionless units), and the streamers propagate in an approximately symmetric manner in both directions of the electric field. In the simulations of Vitello et al. [6], the background ionization is $10^3/\mathrm{cm}^3$ ($= 2 \cdot 10^{-12}$ in dimensionless units). This low background ionization makes the density gradients much steeper, too steep for the low spatial resolution of the earlier simulations [13]. The improved simulations [6] give a detailed picture of the propagation of negative streamers in low background ionization. The question whether positive streamers do propagate under these conditions, could numerically still not be resolved due to the much steeper gradients. Only the analysis [2] to be summarized below, could explain the nature of the gradients and the conditions of propagation of these fronts.

Thermal ionization of N_2 at room temperature certainly can be neglected, since the Boltzmann factor is about 10^{-100}. The background ionization in the atmosphere is caused by radioactivity and cosmic radiation. An average value is $10^3/\mathrm{cm}^3$ [14], so Vitello's simulations give a realistic picture of a negative streamer discharge in nitrogen – if the use of a continuum approximation for the initially very low density $10^3/\mathrm{cm}^3$ is appropriate. However, it seems reasonable to assume that during the evolution times of the simulations this small initial density is effectively negligible. Then these simulations show a picture of an essentially non–ionized outer area penetrated by the streamer. We will discuss the situation of streamers penetrating a non–ionized area in more detail below.

The simulations of Dhali and Williams [13] with the higher background ionization are realistic either for a pre–ionized gas or for semiconductor systems that actually can be thermally ionized at room temperature due to much lower ionization energies. A qualitative theory for such streamer discharges actually has been developed by D'yakonov and Kachorovskii [5]. Here the streamer velocity is essentially determined by the size of the "active" area around the streamer tip where the electric field is above the ionization threshold. The ionization wave then is a collective mode based on impact ionization and electric screening, which can become much faster than the speed of the local charge carriers. This is why the currents of the charge carriers are neglected in [5]. Therefore two basically different propagation modes should be distinguished: streamers in negligible background ionization and streamers in pre–ionized media.

Streamer–like impact ionization fronts propagating into a pre–ionized medium also are the mode of operation of semiconductor devices such as the microwave TRAPATT–diode [15] and the pulse sharpener diode [16]. In these layered semiconductor structures, depletion layers, small electrode distances or load resistors

in the external circuit lead to additional stabilization or destabilization mechanisms [17,18] of planar impact ionization fronts.

2.3 Planar Streamer Fronts

We now recall the basic analytical properties of planar streamers fronts [2], since this elucidates the dependence on the initial degree of ionization, and also sets the basis for the following analytical steps. Planar fronts that propagate according to the minimal model (5)–(7) with uniform velocity v into an area with little ionization and time–independent homogeneous electric field E^+, solve the following set of nonlinear ordinary differential equations in a comoving coordinate frame $\xi = x - vt$:

$$D\partial_\xi^2\sigma + (v+E)\partial_\xi\sigma + \sigma\partial_\xi E + \sigma f(|E|) = 0 \ , \tag{8}$$

$$D\partial_\xi\sigma - v\partial_\xi E + \sigma E = 0 \ . \tag{9}$$

In the area behind the front, the electric field is screened and the degree of ionization σ^- of the plasma is fully determined by the dynamics, not by any equilibrium process. The fronts can be constructed in the 3–dimensional phase space $(\sigma, \partial_\xi\sigma, E)$ as heteroclinic orbits between fixed points $(0, 0, E^+)$ and $(\sigma^-, 0, 0)$ [2] – where actually the whole E– and σ–axes are fixed point lines. In contrast to the more common fronts between a stable and a metastable state, the velocity v here is not uniquely determined by the far field E^+. Rather a negative front can have any velocity $v \geq v^*$, where

$$v^* = |E^+| + 2\sqrt{D\ f(|E^+|)} \ , \tag{10}$$

and a positive front can have any velocity $v \geq v^\dagger$, where

$$v^\dagger = D\ \tilde{v}^\dagger(E^+) + O(D^2) \ . \tag{11}$$

$\tilde{v}^\dagger(E^+)$ is the nonlinear eigenvalue of a strongly heteroclinic orbit of eqs. (8)–(9) after rescaling with D and then letting $D \to 0$, for details see [2]. Here negative/positive fronts are defined as carrying a negative/positive space charge in the front region and propagating parallel/antiparallel to the drift direction of electrons.

There is a continuous family of front solutions for given E^+, because the penetrated non–ionized state $(0, 0, E^+)$ is linearly unstable with respect to an infinitesimally small electron density σ [2,19]. If the electron density $\sigma(x, 0)$ initially vanishes as $e^{-\lambda x}$ for large x with some $\lambda < \lambda_{steep}$, where

$$\lambda_{steep} = \begin{cases} \sqrt{f(|E^+|)/D} & \text{neg. fronts,} \\ \left(v^\dagger + E^+ - \sqrt{(v^\dagger + E^+)^2 - 4D\ f(|E^+|)}\right)/(2D) & \text{pos. fronts,} \end{cases} \tag{12}$$

it actually will conserve its spatial decay rate in time and will approach an asymptotic front attractor with velocity $v = -E^+ + D\lambda + f(|E^+|)/\lambda$ larger

than v^* or $v^\dagger$. If the spatial decay of σ initially is slower than exponential, a planar front will accelerate. The slower the spatial decay, the higher will be the speed, and the more positive and negative fronts will resemble each other. This observation illustrates the similar propagation of positive and negative fronts in a system with high initial ionization [13].

On the other hand, sufficiently steep initial ionization profiles with

$$\lim_{x\to\infty} \sigma(x,0)\, e^{\lambda_{steep} x} = 0 \tag{13}$$

for large times approach the "selected" velocity v^* or $v^\dagger$ [2,19]. In the case of negligible initial ionization, the slow propagation of positive fronts with velocity $v^\dagger \propto D$ and with a degree of ionization $\sigma^- \propto 1/D$ behind the front [2] results from the balance of electron drift and diffusion, which leads to a singular dependence on D. This explains the numerical difficulties of the simulations of positive fronts [6]. However, the neglect of the ion motion then ceases to be a valid aproximation. Hence we will not consider this solution any further. The negative front with sufficiently low background ionization, in contrast, propagates with the velocity v^* (10) that is slightly higher than the drift velocity $|E^+|$ of an electron in an electric field E^+. The limit $D \to 0$ is smooth in this case [2].

3 Streamers as a Laplacian Growth Problem

In the remainder of the paper, we discuss the propagation of negative streamer fronts into an initial state with negligible ionization, and we will refer to the simulations [6] of Fig. 1.

For numerical convenience, the initial conditions of the simulations [13,6] have a considerable ionization in some bounded part of space. Since such initial conditions are somewhat arbitrary and also might differ depending on prehistory, Vitello *et al.* [6] study in detail various initial conditions. It turns out that the late stage development of the propagating tip of the finger is always quite the same. This motivates the working hypothesis, that a certain late stage finger–shape is an attractor of the dynamics for a large class of initial conditions.

3.1 Pattern Formation on the Outer Scale: Laplacian Growth

That finger–shaped moving interfaces can be attractors of the dynamics, is known from other pattern forming systems like viscous fingers or dendrites in solidification fronts [20]. This analogy can be made more precise [2]: The smallest length scale in Fig. 1 [6] is the width of the transition region between non–ionized and ionized region. If this length scale can be eliminated, the front can be modeled by an effectively infinitely thin interface on the outer scale.

Then the following reduced model emerges: The ionized area is conducting, so the field is low and the electric potential is

$$\Phi \approx \text{const.} \quad \text{in the ionized region.} \tag{14}$$

In the non–ionized area, there are no space charges, so that

$$\nabla^2 \Phi = 0 \quad \text{in the non–ionized region.} \tag{15}$$

If the electric field immediately in front of the ionization front has the time–independent value $\mathbf{E}^+$, then a planar front approaches a stationary state with velocity (10)

$$v^* = |\mathbf{E}^+| + 2\sqrt{Df(|\mathbf{E}^+|)} \quad , \quad |\mathbf{E}^+| = \hat{n} \cdot \nabla \Phi|_+ \ . \tag{16}$$

This is the velocity of the interface or moving boundary in the outer scale description. Here $\hat{n}$ is the orthonormal vector on the interface, and $\nabla \Phi|_+$ denotes the extrapolation of $\nabla \Phi$ from the non–ionized region onto the interface.

The moving boundary problem defined by (14)–(16) is the same as the lowest order aproximation of viscous fingering in two dimensions, where an incompressible fluid of negligible viscosity penetrates an incompressible fluid of finite viscosity in a Hele–Shaw cell. The potential Φ then is replaced by pressure p.

It has been established in the gas discharge literature [21], that the model (14)–(16) with appropriate outer boundary conditions has uniformly translating paraboloid solutions. Mathematically speaking, these are equivalent to the Ivantsov paraboloids in dendritic growth [20].

3.2 The Need of a Microscopic Stabilization Mechanism

However, from studies of viscous fingering and dendritic growth, it is also known, that all paraboloid solutions are dynamically unstable under the dynamics of (14)–(16). In fact, an arbitrary initial condition generically will develop singularily curved interfaces within finite time. This fact is related to the short wave length instability of linear perturbations of planar fronts. A transversal Fourier perturbation e^{ikx+st} of a planar ionization front that propagates into a non–ionized area with field E^+, has the dispersion relation [22]

$$s(k) = c(E^+)\, k \quad , \quad c(E^+) = \left. \frac{dv^*(E)}{d \ln |E|} \right|_{E^+} \ . \tag{17}$$

The only difference to viscous fingering or dendritic growth is the weakly nonlinear dependence of v^* on E^+ for $D \neq 0$, which causes $v^*(E^+)$ and $c(E^+)$ to differ slightly.

In viscous fingering, e.g., an additional length scale not included in the model (14)–(16) is set by surface tension [20]. Inclusion of this length scale into the problem changes the dispersion relation to the form $s(k) = ck - dk^3$ and removes the short wave length instability. It also allows the accurate prediction of the actually realized finger–shape. The same holds for dendritic growth. Hence the question rises, which microscopic scale stabilizes the approximately uniformly translating shape of the streamer finger in Fig. 1.

3.3 Moving Boundary Approximations and Pulled Fronts

The separation of inner and outer length scales that can be seen in the simulations, suggests the development of a moving boundary approximation from the streamer equations (5)–(7). Such an approximation substantiates the interface approximation above by matching different asymptotic expansions for the inner and outer scale. In particular, on the outer scale description of (14)–(15), it will lead, e.g., to a curvature correction to the interface velocity (16). Hence it seems appropriate to translate the derivation [23,24] of a moving boundary approximation for PDE problems to the streamer equations in order to derive the corrections to (16) in a systematic manner.

A direct translation of these calculations to the streamer model, however, leads to a failure that can be placed in a broader context: the negative streamer front propagating into a non–ionized state belongs to the larger class of so–called pulled fronts [19]. For pulled fronts, the standard calculational methods break down. As is explained in more detail in [25], the calculations in [23,24] assume implicitly that the interface can be parametrized by its geometric properties like curvature. The corrections to the interface velocity then are evaluated by solvability integrals, that arise from multiplication of the equations in first order perturbation theory of the form $\mathcal{L}\phi_1(\xi) = g(\xi)$ with a left zero mode of the linear operator $\mathcal{L}$. If the zero order solution about which one linearizes, is a pulled front, the solvability integrals diverge. This divergence is related to the divergence of the dynamically dominant area: while for fronts propagating into a metastable state or for so-called pushed fronts – an example of a pushed front is the positive streamer front propagating with velocity $v^\dagger$ (11) –, the dynamics is determined in the interior nonlinear area of the front, for pulled fronts the dynamics is determined in the leading edge of the front, where linearization about the non–ionized state is a valid approximation. This area is half–infinite, and hence the spatial integral diverges.

This divergency can not be removed by an appropriate regularization, since it signals the break–down of the underlying ansatz. The ansatz as in [23,24] explicitly uses a separation of length scales between the width of the front and the outer scales, but it implicitly also assumes a separation of time scales between inner and outer dynamics. If the front is described by its geometric properties like curvature only, an adiabatic elimination of the inner time scale is assumed. This is not possible, if there is no inner time scale, and this is exactly the case for pulled fronts. This lack of an inner time scale is illustrated in particular by the relaxation of a negative planar streamer front with sufficiently steep initial conditions (13) towards its asymptotic velocity v^* (10). It is asymptotically for large t

$$v(t) = |E^+| + \sqrt{Df(E^+)} \left[2 - \frac{3}{2\,\tau} + \frac{3\,\sqrt{\pi}}{2\,\tau^{3/2}} + \mathcal{O}\left(\frac{1}{\tau^2}\right) \right] \ , \ \tau = f(E^+)\,t\ . \quad (18)$$

This algebraic relaxation that is universal for all pulled fronts [19], has up to order $1/t$ no intrinsic time scale, in contrast to the more common pushed or bistable fronts that relax exponentially fast in time.

3.4 Possible Stabilization Mechanisms

The problems in the formal derivation of a moving boundary approximation from the streamer equations (5)–(7) are due to the leading edge dominated dynamics of pulled fronts. The exponentially decaying leading edge is generated by diffusion: for $D = 0$, the ionization front in a field E^+ becomes a shock front propagating with the drift velocity $v = |E^+|$. Inspection of (18) shows, that for $D = 0$ also the algebraic relaxation is removed. Still there is a length scale in the problem, which is not yet incorporated in the lowest order interface approximation (14)–(16), namely the electric screening length $\ell_\alpha = 1/\alpha(E^+) = |E^+|/f(E^+)$. It is related to the Maxwell relaxation time $t_M = 1/f(E^+)$.

The planar uniformly translating negative streamer shock front for $D = 0$ can be solved analytically: assume that in the comoving frame $\xi = x - vt$, the position of the ionization shock is $\xi = 0$. For $\xi \geq 0$, there are no charge carriers, and the electric field has the constant value $E^+ < 0$. For $\xi < 0$, one can express electron and ion densities σ and ρ and coordinate ξ as a function of E as [2]

$$\rho[E] = \int_{|E|}^{v} d\epsilon\, \alpha(\epsilon) \quad , \quad \sigma[E] = \frac{v\, \rho[E]}{v + E} \quad , \quad \xi = \int_{|E(-\xi)|}^{v} d\epsilon\, \frac{v - \epsilon}{\epsilon\, \rho[\epsilon]} \quad , \quad v = |E^+| \; . \tag{19}$$

At $\xi = 0$, σ jumps to the finite value $\sigma(0^-) = f(E^+) = |E^+|\, \alpha(E^+)$, and then decreases for $\xi \to -\infty$ to the neutral plasma density $\sigma(-\infty) = \int_0^v d\epsilon\, \alpha(\epsilon)$. Also the derivatives of ρ and E are discontinuous at $\xi = 0$.

Can the additional length scale ℓ_α lead to a stabilization of the short wave length instability of (17)? Preliminary results [22] do show the appearance of the length scale $\alpha(E^+)$ in the dispersion relation

$$s(k) = \begin{cases} |E^+|\, k & \text{for } k \ll 1 \\ |E^+|\, \alpha(E^+)/2 & \text{for } k \gg 1 \end{cases} \; , \tag{20}$$

but a short wave length instability still stays present in the dispersion relation of linear transversal perturbations of a planar front.

At present, we investigate two working hypotheses: either diffusion D needs to be included to remove the short wave length instability, or the instability is removed from a sufficiently curved drift front with $D = 0$ by electrostatic effects [22]. This would then allow an outer scale analysis of improved interface equations (14)–(16) and a prediction of the asymptotic form of a streamer finger.

4 Summary and Outlook

We have discussed two pattern forming systems in electric discharges, namely barrier and streamer discharges. In barrier discharges, one can realize, e.g., stationary patterns in layered systems with wide lateral aspect ratios, which show similarities to, e.g., Rayleigh–Benard–convection. Streamer discharges are transient finger–shaped ionized patterns penetrating an area with very low or vanishing ionization, similar to viscous fingers or dendrites.

Both barrier and streamer discharges are characterized by low ionization and current densities. Hence the patterns in simple media like pure nitrogen should be described by the minimal discharge model (5)–(7). This model is a reaction–drift–diffusion model for two charged species coupled to the electric field. The field determines reaction and drift, while the space charges modify the field. This leads to nonlinear dynamics with a ionization–field–coupling, if the particle densities and hence the space charges are sufficiently large.

In particular, we discussed the penetration of a non–ionized area with high electric field by an ionized negative streamer finger. We identified this problem as a Laplacian growth problem quite comparable to viscous fingering or dendritic growth. For a complete moving interface approximation of streamer propagation, a correction to the velocity due to curvature or relaxation is still missing. This is due to the particular properties of negative streamer fronts that belong to the larger class of pulled fronts. For pulled fronts, standard moving boundary approximations fail, because there is no separation of inner and outer time scales. We presented first steps towards the derivation of a different approximation.

Acknowledgement

We thank P. Rodin for productive discussions about streamer propagation in pre–ionized media, and for advice and references on streamer discharges in semiconductor devices. Many aspects of the nonlinear analysis presented in this review were shaped during previous collaborations of U.E. with W. van Saarloos. We thank P.A. Vitello for the permission to reprint Fig. 1. U.E. was supported by the Dutch Science Foundation NWO, and M.A. by the EU-network "Patterns, Noise, and Chaos".

References

1. Y.P. Raizer: *Gas Discharge Physics* (Springer, Berlin 1991)
2. U. Ebert, W. van Saarloos, C. Caroli: Phys. Rev. Lett. **77**, 4178 (1996); and Phys. Rev. E **55**, 1530 (1997)
3. J. Meunier et al.: J. Appl. Phys. **78**, 731 (1995); C. Punset et al.: J. Appl. Phys. **83**, 1884 (1998)
4. E. Schöll: *Nonequilibrium Phase Transitions in Semiconductors* (Springer, Berlin 1987)
5. M.I. D'yakonov, V.Y. Kachorovskii: Sov. Phys. JETP **67**, 1049 (1988); and Sov. Phys. JETP **68**, 1070 (1989)
6. P.A. Vitello, B.M. Penetrante, J.N. Bardsley: Phys. Rev. E **49**, 5574 (1994)
7. Yu. Astrov, E. Ammelt, S. Teperick, H.-G. Purwins: Phys. Lett. A **211**, 184 (1996); Yu. Astrov et al.: Phys. Rev. Lett. **78**, 3129 (1997), Phys. Rev. Lett. **79**, 2983 (1997), Phys. Rev. Lett. **80**, 5341 (1998); E. Ammelt et al.: Phys. Rev. E **55**, 6731 (1997), Phys. Rev. E **58**, 7109 (1998); C. Strümpel et al.: Phys. Rev. E **61**, 4899 (2000)
8. E. Ammelt, D. Schweng, H.-G. Purwins: Phys. Lett. A **179**, 348 (1993); I. Brauer et al.: J. Appl. Phys. **85**, 7569 (1999), Phys. Rev. Lett. **84**, 4104 (2000); I. Müller et al.: Phys. Rev. Lett. **82**, 3428 (1999), IEEE Trans. Plasma Science **27**, 20 (1999)

9. C. Radehaus et al.: Phys. Rev. A **45**, 2546 (1992)
10. Yu. Astrov: [private communication]; D. Sijacic, U. Ebert: [in preparation]
11. W. Breazeal, K.M. Flynn, E.G. Gwinn, Phys. Rev. E **52**, 1503 (1995)
12. H. Raether: Z. Phys. **112**, 464 (1939) (in German); and Ergebnisse der exakten Naturwissenschaften **22**, 73 (1947) (in German)
13. S.K. Dhali, P.F. Williams: Phys. Rev. A **31**, 1219 (1985); and J. Appl. Phys. **62**, 4696 (1987)
14. A.M. Howatson: *An Introduction to Gas Discharges*, 2nd edn. (Pergamon, Oxford 1976)
15. B.C. Deloach, D.L. Scharfetter: IEEE Trans. El. Dev. **ED-20**, 9 (1970)
16. I.V. Grekhov, A.F. Kardo-Sysoev: Sov. Tech. Phys. Lett. **5**, 395 (1979)
17. S.N. Vainstein, Yu.V. Zilyaev, M.E. Levinstein: Sov. Tech. Phys. Lett. **14**, 920 (1998); Sov. Phys. Semicond. **22**. 717 (1988)
18. A.M. Minarsky, P.B. Rodin: Semiconductors **31**, 356 (1997); Sol. St. Electron., **41**, 813 (1997)
19. W. van Saarloos: Phys. Rev. A **39**, 6367 (1989); U. Ebert, W. van Saarloos: Phys. Rev. Lett. **80**, 1650 (1998); and Physica D (2000) [in press]
20. For a review and a collection of older relevant articles see P. Pelcé: *Dynamics of Curved Fronts* (Academic Press, San Diego 1988)
21. This can be concluded from a remark in [5], referring to E.D. Lozanskii, O.B. Firsov: *Spark Theory* (Atomizdat, 1975) (in Russian). From the remark, the solution can be reconstructed using the calculation of the field outside any conducting ellipsoid, see, e.g., L.D. Landau, E.M. Lifshitz, paragraph 4 of vol. VIII: *Electrodynamics of Continuous Media* (Pergamon, Oxford 1960)
22. M. Arrayás, U. Ebert [in preparation]
23. P. Fife: *Dynamics of Internal Layers and Diffusive Interfaces* (SIAM, Philadelphia 1988)
24. A. Karma, W.J. Rappel: Phys. Rev. E **53**, R3017 (1996); and Phys. Rev. E **57**, 4323 (1998)
25. U. Ebert, W. van Saarloos: Phys. Rep. (2000) [in press]
(1988) Thesis, (1977)

Evidence for Eigenfrequencies in Dendritic Growth Dynamics

Jeffrey C. LaCombe, Matthew B. Koss, Cindie Giummarra, Julie E. Frei, Afina O. Lupulescu, and Martin E. Glicksman

Materials Science and Engineering Department, Rensselaer Polytechnic Institute, Troy, New York 12180-3590, USA
E-mail: lacomj@rpi.edu

Abstract. Microgravity dendritic growth experiments, conducted aboard the space shuttle Columbia, are described. In-situ video images reveal that pivalic acid dendrites growing in the diffusion-controlled environment of low-earth orbit exhibit a range of transient or non-steady-state behaviors. The observed transient features of the growth process are being studied with the objective of understanding the mechanisms responsible for these behaviors. Included in these observations is possible evidence for characteristic frequencies or limit cycles in the growth behavior near the tip of the dendrites. These data, and their interpretations, will be discussed.

1 Introduction

The formation of dendritic microstructures is common during the solidification of most metals and alloys. It is important to understand the process by which these dendrites form because their microstructures can persist through subsequent material processing stages, and affect the properties of the finished product.

Over the past five decades, the scientific community has produced a large body of theoretical and experimental work describing dendritic growth. Glicksman and Marsh [1] discuss this research in their 1993 review article, and Bisang and Bilgram [2] also include a detailed review as part of their 1996 article on the subject. As these reviews and others make clear, the speed of an advancing dendrite tip is a critical characteristic of dendritic growth. In 1947, Ivantsov [3] introduced two simplifying assumptions in the theory describing thermal dendrites growing in a supercooled melt: 1) that a dendrite can be represented as a shape-preserving paraboloidal interface with a tip radius, R, 2) that the dendrite grows at a constant rate, V. The notion that real dendrites (as opposed to paraboloidal needle-crystals) also grow in a steady-state manner, with constant velocity, is generally considered to be supported by numerous experimental observations of isolated dendrites (see Huang for example [4]). Most experimental and simulation studies of dendritic growth thus attempt to extract a constant velocity measurement as a parameterization of the kinetics, whereas most theoretical studies assume constant velocity behavior.

Recent work has focused on studying the thermal interactions between a dendrite and its surroundings, and quantifying what is required for a dendrite to grow in a truly "isolated" manner. Thermal interactions may exist between

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 283–297, 2001.
© Springer-Verlag Berlin Heidelberg 2001

a dendrite and its neighboring tips, container walls [5], or even its own initial structure [6] and trailing side branches [7, 8]. Such interactions are potential explanations for experimentally observed growth rates that do not match the predictions for an isolated diffusion-limited dendrite, with the thermal boundary conditions set at infinity. If the strength of these thermal interactions changes over time, the velocity of the dendrite should change as well, and in general, there should be deviations from the conventionally prescribed behavior predicted by an Ivantsov-like solution for an isolated dendrite. In particular, at low supercoolings, the length scale over which the thermal field extends is known to be large compared to the morphological length scales relevant to the dendritic growth process. With such long thermal diffusion lengths, proximate dendrite arms can mutually interact to generate an operating state that is not steady. This type of behavior is evident in the recent phase-field simulations by Provatas *et al.* [9], which suggest the existence of an early-time transient in two-dimensional dendritic growth prior to approaching a steady-state growth rate.

2 Experiment

The data presented here derives from an analysis of dendritic growth data obtained from the Isothermal Dendritic Growth Experiment (IDGE). This project compiled data on the growth of pivalic acid (PVA) dendrites in microgravity. Additional detailed information describing earlier IDGE experiments can be found elsewhere [10].

2.1 Apparatus

The experimental apparatus used for the IDGE experiments was located in the payload bay of the space shuttle *Columbia*. This equipment, depicted in the schematic of Fig. 1, was specially designed for use on earth as well as the environment of low-earth orbit. The sample material was contained by a rigid quartz growth chamber located within a temperature-controlled bath. The growth chamber interior volume measured approximately 31mm square by 50 mm long. Nucleation was achieved through the use of a hollow stinger tube that penetrated the wall of the growth chamber. The exterior end of the stinger tube was closed and surrounded by a thermoelectric cooler. The interior end was open, allowing the PVA sample material in the chamber to also fill the stinger.

The use of a quartz growth chamber was necessitated by the sample material. Pivalic acid was found to be excessively reactive with the stainless steel and borosilicate glass growth chambers used in earlier IDGE space flight experiments where the sample material was succinonitrile [10]. The PVA filled stainless steel chambers rapidly degraded in purity, and would not supercool enough to conduct the necessary experiments. Additionally, for reasons that are not completely understood, the PVA-filled stainless steel chambers were also prone to structural failure.

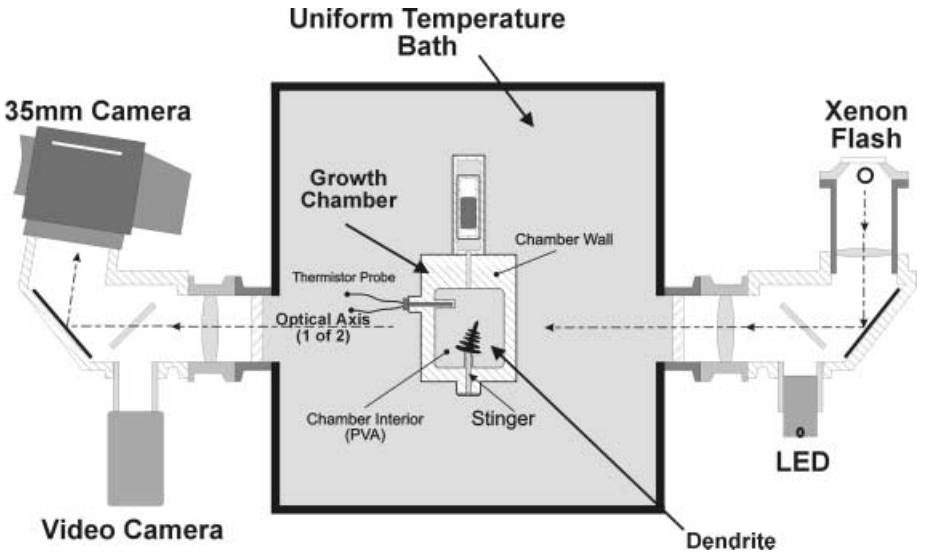

Fig. 1. Schematic representation of the experimental apparatus used in these experiments. The sample material is located inside the quartz growth chamber. This is placed within a temperature-controlled bath where the experiment is conducted with $in-situ$ monitoring of process parameters (temperature *etc.*) and imaging takes place.

During the operation of the experiment, each dendritic growth cycle began by completely melting the PVA, followed by lowering the melt's temperature to the desired supercooling. After the supercooled melt's temperature reached steady-state, the thermoelectric cooler was activated. This nucleated a small crystal in the end of the stinger, which then propagated down the stinger tube to emerge into the chamber as a freely growing dendrite. Once one of these "growth cycles" was completed, a new growth cycle was initiated by re-melting the sample and proceeding as described above. This arrangement, combined with the microgravity conditions, produced dendritic crystals grown under diffusion-limited conditions, with the bath temperature controlled to within 0.002 K (spatially and temporally). The same apparatus was also operated on the ground to acquire an appropriate "baseline" data set describing dendritic growth under the convective conditions associated with earth-gravity.

During the growth cycles, once a crystal emerged from the stinger, images of the dendrites were obtained from two perpendicular views using both electronic and film cameras. While the film images constituted the primary data source for the IDGE experiments, the observations described here are derived primarily from the electronic video cameras (Fig. 2). These cameras provided the spatial and temporal resolution that is necessary to study the transient aspects of the growth process. Specifically, an imaging chip array was used of 640 x 480 pixels (256 gray-scale levels) and an imaging rate of 29.963 frames per second ($\sim$30 Hz).

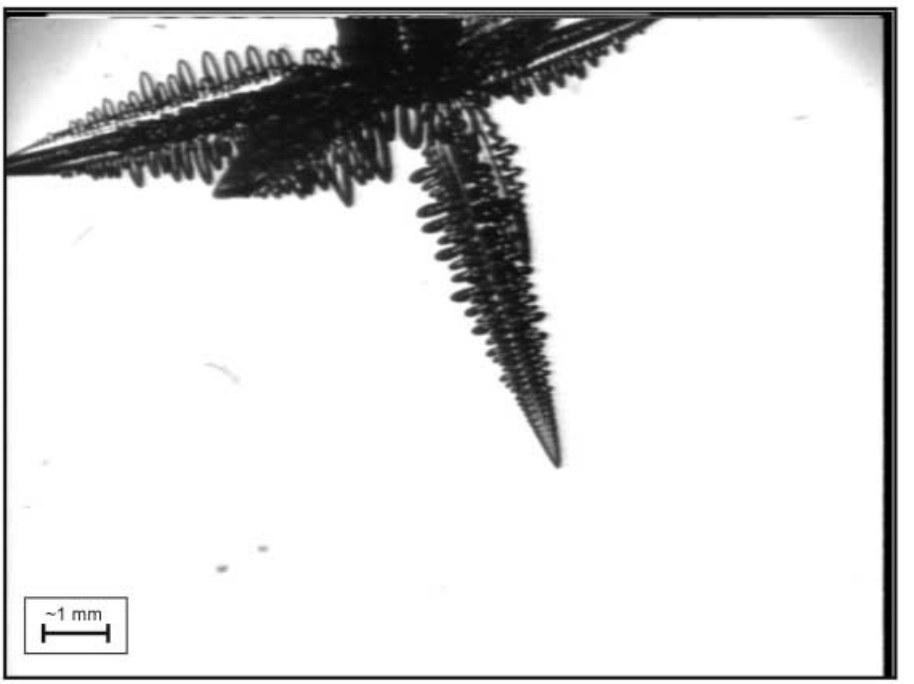

Fig. 2. Video image of a dendrite from Cycle 04 grown at 0.407 K supercooling during the USMP-4 experiment. This and other similar images were used to produce the results presented here.

2.2 Data Extraction

The data presented here are primarily based upon measurements of dendrite tip locations as a function of time, obtained from the 30fps video cameras. The base optical resolution of the system is related to the size of an individual pixel in the video camera's imaging array. Each pixel, after correcting for the magnification of the optics system, images a region of the chamber that is approximately 22 μm by and 22 μm. These values also constitute the raw measurement precision for the tip position data. It is beneficial to improve upon this precision by applying a sub-pixel resolution image analysis method to each image in the growth cycle (see Figs. 2 and 3 for a sample images).

The method used for the data presented here begins by first examining the first and last image frames of a growth cycle. The dendrite tips are found in these images using a row-by-row search to locate the lowest pixel in the field of view that is darker than a specified threshold value. If there are more than one pixel in a row satisfying this criteria, an average is calculated to obtain the horizontal coordinate. With these crude estimates of the tip positions at the beginning and end of the growth, it is possible to construct a vector used to predict where the tip will be in all frames during the growth, provided the frame number and frame

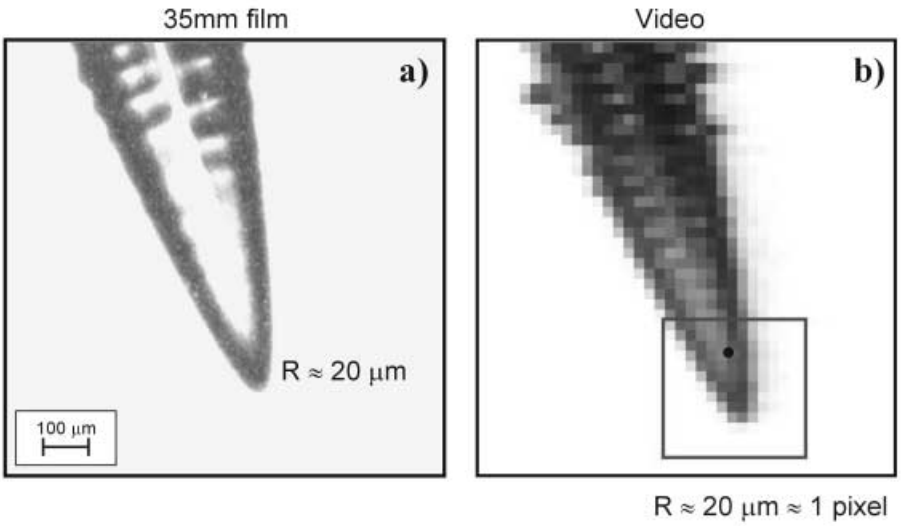

Fig. 3. Images of the tip region of dendrite grown in Cycle 04. (a) A 35mm film image with approximate tip radius of curvature 20μm. (b) 30fps video image, revealing considerably less detail. The centroid of the tip is calculated using the pixel intensities located within a box enclosing the tip.

rate are incorporated. This tip-location prediction scheme improves processing efficiency, though is only capable of resolving information on the order of 22 microns (one pixel).

Using this predicted tip location, a second stage of refinement is added to the tip-location process for each image. This is achieved by overlaying a "sampling line" along the predicted vector and determining the point along this vector where the image intensity crosses a selected threshold value. However, this time the sampling line is "thick" in that it also comprises of several pixels on either side of the mathematical line, creating an averaging effect. Additionally, interpolation is applied to determine the threshold location more precisely. By incorporating a statistically larger number of pixels in this second stage tip locating method, resolution is improved to approximately 7 microns ($\sim$1/3 pixel). However, the horizontal coordinate is constrained to fall along the predicted vector line. In practice, this is not usually a significant issue, since these dendrites tended to grow within about 10 degrees of vertical, and, once started, do not deviate in direction under nominal growth conditions.

The final stage of resolution enhancement is achieved by a somewhat unorthodox approach, which has the advantage of incorporating still more (statistically speaking) information concerning the tip location. Fig. 3b shows a typical video image of a dendrite tip. Using the refined tip location (stage 2, described above) as a reference location, a box is created around the tip. Next, the centroid of the pixel intensity within this box is calculated. In practice, the coordinates of

this centroid exhibit a resolution of approximately 2 microns (∼1/10 pixel). This approach doesn't actually locate the tip's interface. Instead, it uses more data to obtain a more consistent reference point, which serves to track the tip's movement over time. The size of the sampling box used in the centroid calculation is somewhat arbitrary. The concerns in its selection are primarily in obtaining a balance between the desire to have a large number of data points contributing to the measurement, and avoiding the inclusion of side branches. When information that is obtained from regions further removed from the tip is used, nascent side branches can contribute to the centroid calculation in a periodic manner.

Once extracted from the images, the tip positions are then converted into a displacement vs. time data set that is used for subsequent analysis (see Fig. 4). The displacement is calculated relative to the tip position in the first available frame of video data, which is assigned to time = 0. The measures of resolution quoted herein are characterized by examining the residuals produced by a linear regression of the displacement data (Fig. 5a). The standard deviation of the spread in these residuals is calculated on a moving basis, with a 2-second window. These measurements are shown in Fig. 5b, revealing the 2 micron (at most) uncertainty stated above.

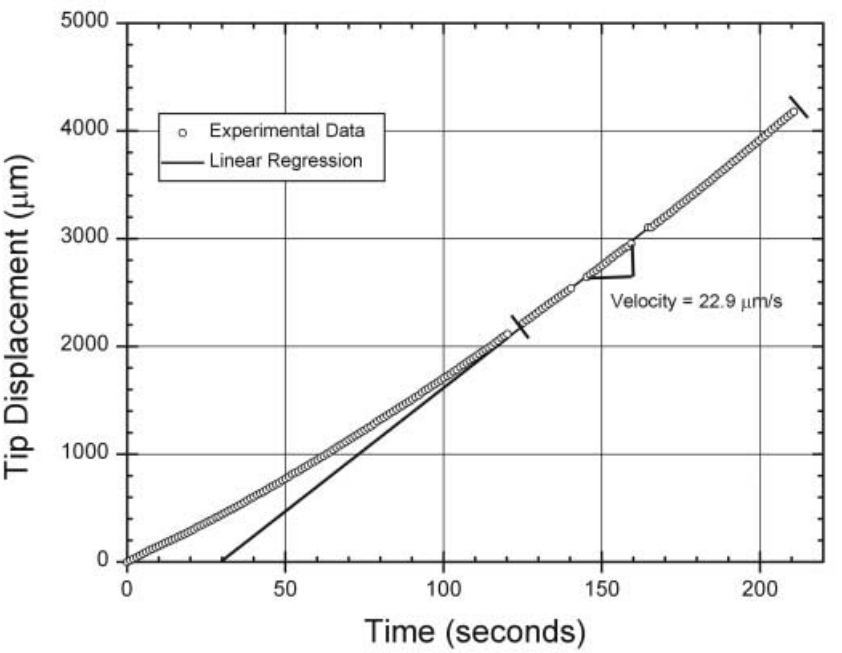

Fig. 4. Tip displacement vs. time for dendrite Cycle 09, grown at 0.376 K supercooling. Linear regression performed using data from the more steady-state portion of the growth (after ∼125 seconds). The growth rate, 22.9 μm/s results from the slope generated by the regression. Note: 1 of every 30 data points plotted.

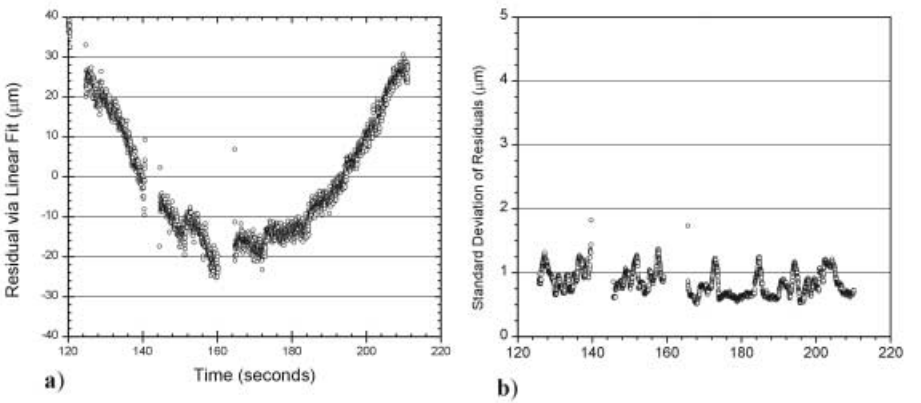

Fig. 5. (a) Residuals resulting from a linear fit to the tip displacement vs. time data in Fig. 4 for the latter growth stage (i.e., after ∼125 seconds). (b) The standard deviation of the spread in the residuals is below ∼2 μm, representing the uncertainty in the tip-position measurements.

2.3 Spectral Analysis

The tip displacement vs. time data, obtained as described above, is evaluated here to identify the possible presence of characteristic frequencies. Commonly, a discrete Fourier transform (DFT) is used in this type of analysis, although they have inherent limitations. In order for established FFT (fast Fourier transform) algorithms to apply, it is necessary that the subjected data be sampled at even spacings in time. Additionally, the FFT algorithms do not handle gaps in the sampled data well, and many have been "tuned" to operate on data sets of size equal to a power of 2 (for computational efficiency). A method for the spectral analysis of unevenly sampled data was developed by Lomb [11] (which is partially based on work by others) and is described well in the popular "Numerical Recipes" programming handbooks [12]. The Lomb approach produces a *normalized periodogram*, with spectral power as a function of frequency. Through the normalization, a *significance level* is also produced, which serves as a measure of the probability that a spectral peak is a "false alarm". The spectral analysis presented here derives from the output of this Lomb method as implemented in a program written in Microsoft Visual Basic© . This effectiveness of this method was confirmed through numerous "test cases", where data sets were created with noise superimposed onto multiple sinusoidal signals. The Lomb method successfully identified the correct frequencies in these test data sets with signal amplitudes as small as $1/15^{th}$ of the maximum (random-number generated) noise amplitude. Additionally, the Lomb method was satisfactorily compared with output from an FFT algorithm. It should be noted that the method used here to reveal characteristic frequencies differs from the one described in Dougherty

et al. [13], which tracked the side branch amplitude on one side of a dendrite over time, at a fixed distance behind the tip.

3 Results

Earlier observations by the authors [14] indicated that dendritic growth rates are not constant over the time-scale of observation. Specifically, the dendrite tip, after completing an initial transient, continues to experience a small acceleration. It is not yet clear whether or not the dendrite approaches a truly constant velocity. The displacement data shown in Fig. 4 derive from a supercooling of 0.376 K. They are qualitatively representative of the entire body of data analyzed to date. The gaps in the data resulted from the hardware necessity to adjust lighting during 35mm camera exposures. As with the data presented in [14], this cycle shows an initial transient clearly accelerating, and then evolves toward a more constant slope regime from which one can extract the velocity.

Despite the relative linearity of the latter part of the displacement plot, a least-square regression of the data from the portion of the growth between the hash-marks of Fig. 4 (after ∼125 seconds) reveals that the dendrite is still accelerating late in the growth. The residuals resulting from the linear regression show a systematic deviation from steady-state growth (Fig. 5a). A dendrite growing at constant velocity would be expected to exhibit random residuals forming a Gaussian distribution centered about zero. Instead, as shown, a distinct non-random residual occurs, with monotonic upward curvature in the graph. As mentioned earlier, the residuals also quantify the uncertainty in the measurements, because their spread is a good measure of the resolution resulting from the sub-pixel interpolation measurement of the tip location.

A convenient measure of the asymptotic behavior of the tip speed, as evidenced by the non-random residual, is the growth rate exponent, κ, calculated by comparing the data to a power law for the displacement proportional to t^{κ}. An exponent of κ=1 indicates linear displacement in time (i.e. constant velocity). Although at first this method may seem less intuitive than simply comparing the instantaneous velocity versus time, the slope of which is the acceleration, the power law helps elucidate the variations from, and approach to, constant velocity (κ=1) behavior.

The measured values of κ (Fig. 6) calculated after first smoothing the displacement data using a two-second moving average, reveals a slightly different trend than seen for the dendrite growth cycle presented in [14]. In the previous report, κ started out large and approached 1 from above. For this case, κ starts out close to 1, and increases slightly (still an accelerating trend). Aside from this initial transient, other dendrites have been seen to behave qualitatively similar in terms of κ.

Figures 7 and 8 show the results of spectral analysis performed on a dendritic growth cycle 04, grown at 0.397 K supercooling. Several statistically significant peaks are present in the data. Figure 9 reveals that the sizes of some of the

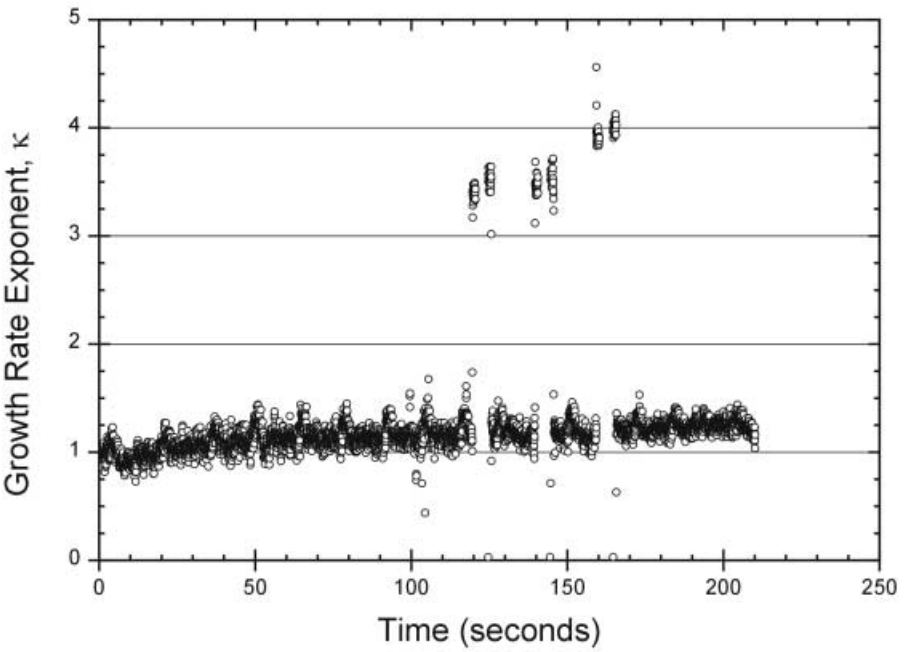

Fig. 6. Growth rate exponent,κ, as a function of time for the dendrite data in Figs. 4 & 5. A value of κ=1 corresponds to steady state, constant velocity growth.

peaks is seen to vary with the amount of the dendrite's tip that is used in the tip-location measurements.

4 Discussion

The data presented here indicates that there are transient aspects of the dendritic growth process which steady-state theory does not describe. Examination of these experimental data suggests that within experimental practice, the establishment of *strict* isothermal conditions at the dendritic interface may be illusory. We will discuss briefly several potential sources for non steady-state behavior. In doing so, it is beneficial to describe the time-dependent process as occurring in two distinct periods. The first of these is the initial transient, which is most evident in Fig. 6 prior to approximately 50 seconds. The second distinct time-dependent period is beyond ∼50 seconds, where the growth rate exponent continues to increase, though more gradually.

To explain the early-growth behavior, the authors [14] suggested two related mechanisms that may induce a transient effect. Summarizing, we note that during the early phase of growth, a dendrite evolves from some pre-dendritic morphological structure to eventually develop an interface shape and thermal field which are commensurate with steady-state growth. This developmental stage

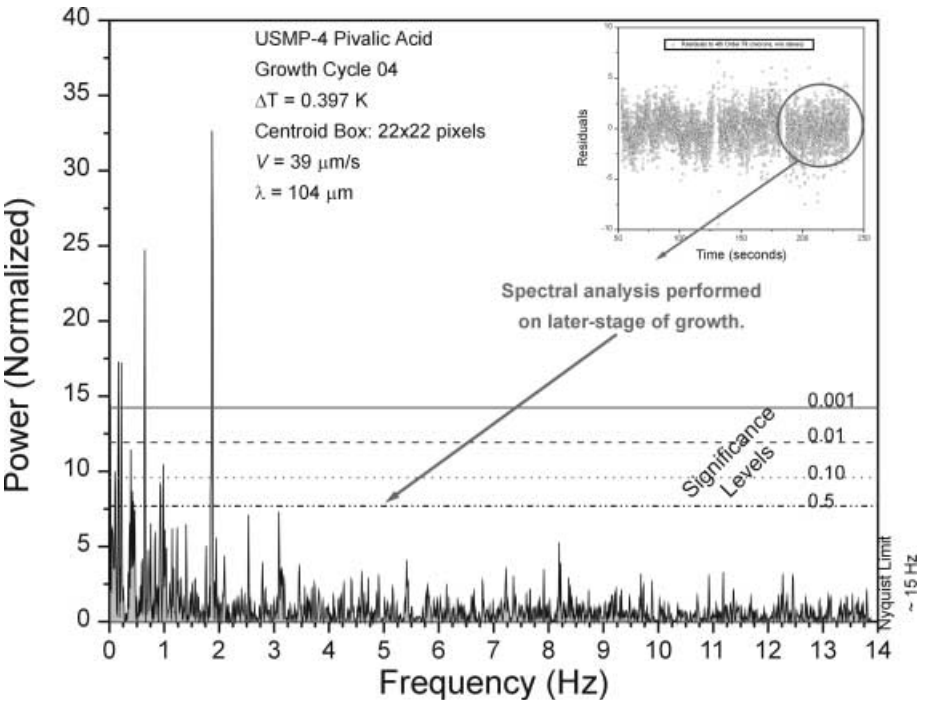

Fig. 7. Spectral Analysis Periodogram. The residuals from a 4^th-order polynomial fit to the displacement data from Cycle 04 were examined using a spectral analysis technique. Supercooling = 0.397 K, Velocity = 39 μm/s, and the side-branch spacing = 104 μm. Also shown are the significance levels, which represent the probability that a peak's power at the specified level is a "false alarm".

takes time. Additionally, in the nascent stages of a dendrite's development, neighboring dendrites (or other branches of an equiaxed dendrite) are located close to one another. As long as these separation distances are less than a few thermal lengths, where the characteristic thermal length is defined here as the thermal diffusivity of the liquid ($7\text{x}10^4$ μm^2/s) divided by the growth rate. Thermal interactions are expected to occur between neighboring dendrites, causing a localized drop in supercooling. As the dendrites grow, their tip-separation increases, creating a progressive increase in the local effective supercooling. This would manifest itself as an accelerating growth rate. Similar interactions also exist for proximate dendrites growing in the same direction. Eventually, if the interactions are strong enough, one of the dendrites will experience favorable thermal conditions and will "pull away" from the slower dendrite.

The experimental (and simulated) observation of an initial transient preceding steady-state growth, is not new. In fact, it has been generally assumed that this phase, although easily distinguishable from the steady-state that emerges, is less germane and fundamental to the kinetics than the steady-state itself. How-

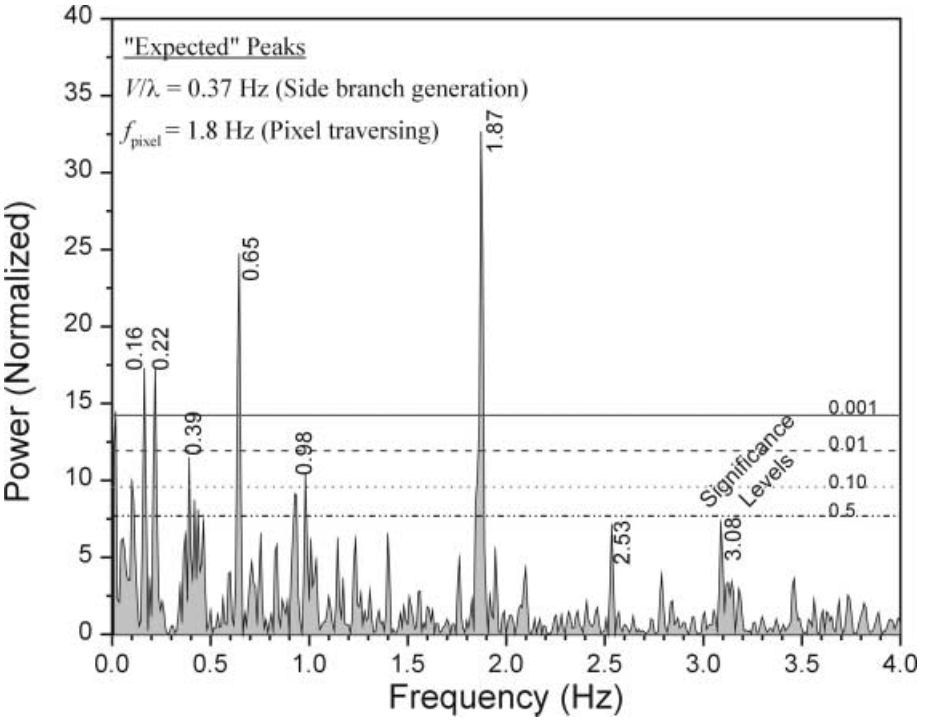

Fig. 8. Detail of Fig. 7. At low frequencies (relative to the Nyquist limit), several statistically significant peaks are featured in the spectrum. Peaks are expected at the frequency at which side arms are being created (0.37 Hz), and at the frequency at which the dendrite tip is traversing across the region of space imaged by each camera pixel (1.8 Hz).

ever, as is revealed here, the growth phase following the initial transient might not ever reach a constant velocity. Thus, the more important issue for such investigations is whether the time-dependent state is fundamental to the physics of isothermal dendritic growth, or appears as an artifact of experiments carried out in finite volumes of supercooled melts.

The observation of the effect of neighboring dendrites on growth velocity [14] suggests possible long-range thermal interactions with proximate dendrites or arms growing in the same general direction as the tip in question. It is conceivable that similar interactions may contribute to a second stage of time-dependent growth seen after the initial transient is completed. However, many of the analyzed dendrites were isolated growths (noting that "isolation" is a strong function of time and supercooling), where the closest neighbor was beyond three thermal lengths. In order to conclusively examine this issue, it will be necessary to analyze a significant number of dendrites growing under a variety of conditions. Additionally, one must firmly establish that the thermal interaction between den-

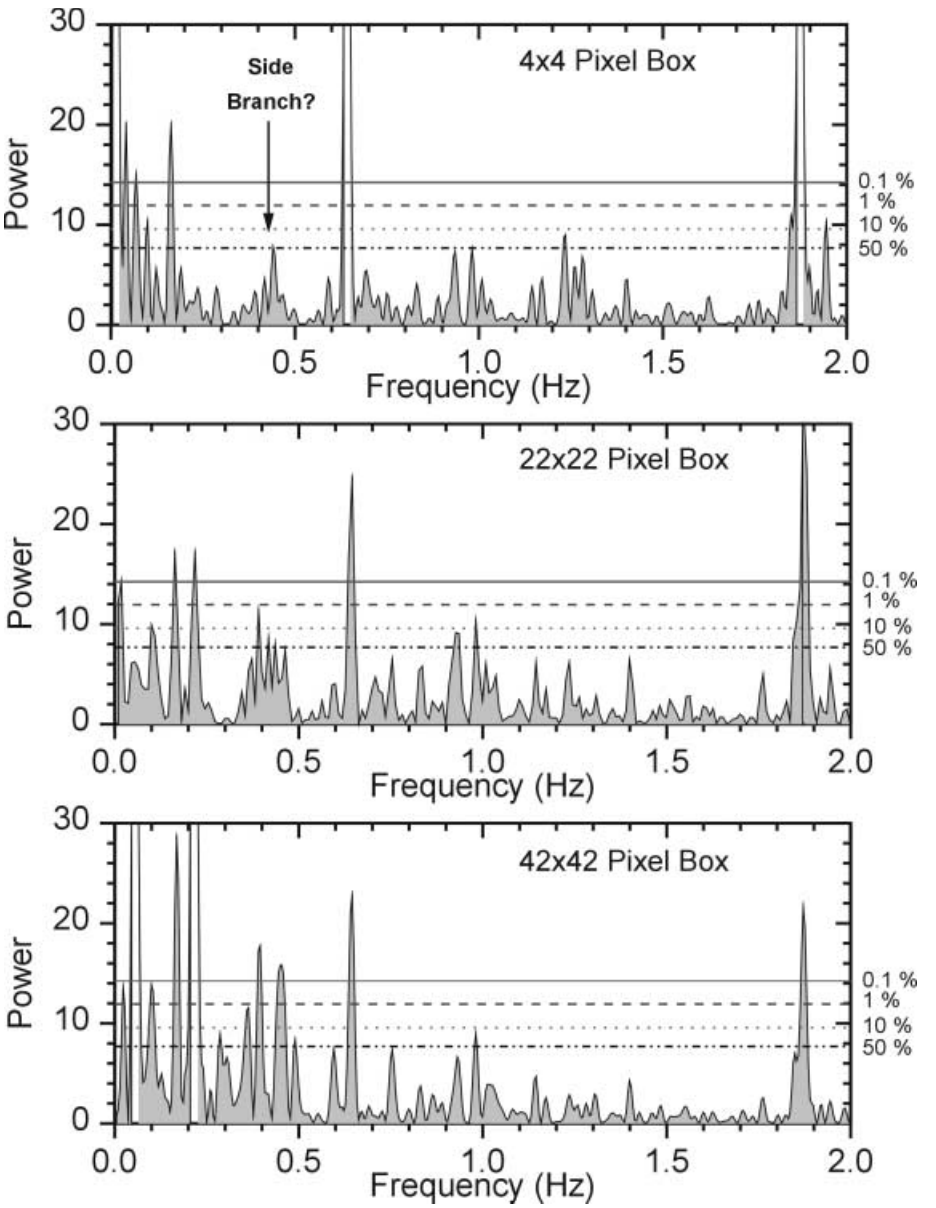

Fig. 9. Cycle 04 was analyzed by locating the tip in each image using a different sized box in the centroid calculation. The resulting displacement data was evaluated as in Figs. 7 and 8. As the sampling box size was increased to include more information from farther behind the tip, the significance of the peak corresponding to the side branching increases. When the dendrite tip is located using data that derives from only the immediate tip region (∼3 Radii), a frequency that is characteristic of the side branches is still present, even though the first detectable side branch (using a microscope) was only observed 10 times further behind the tip.

drites separated by more than three thermal lengths are still enough to produce the measured changes in growth velocity.

The second transient phase, also described in [14], can also potentially be explained by the finite size of the growth chamber. Pines, Chait, and Zlatkowski [5], showed that a dendrite may interact with the walls of a container as growth proceeds. However, it remains to be determined whether the closing rate between the dendrite and the wall is commensurate with the measured time variations in velocity. Given these considerations, it is our hypothesis that it is more likely that the second stage of non steady-state behavior is fundamental to isothermal dendritic growth. Since dendrites are not truly parabolic bodies of revolution [4, 15], there is no compelling phenomenological reason that dendrites *should* grow strictly at a constant rate. As the data set from which these observations and conclusions were drawn is the only one we know of that is both diffusion-controlled, and measured at the necessary temporal resolution to make the described measurements, we plan additional work in data reduction, analysis, and modeling, to explore this hypothesis further.

Spectral analysis identified several statistically significant peaks. The peak seen at 1.87 Hz (Fig. 8), is likely due to the instrumentation used to image the growth process. The electronic camera used in the optical system imaged a region of space approximately 22 μm. A dendrite growing at a speed of 39 μm/s will traverse the rows of pixels at a frequency of 1.8 Hz, revealing the optical resolution limit of $\sim$1 pixel. Some of the other strong peaks may be due to experimental influences related to periodic motion in either the melt (less likely in the microgravity environment) or in the mechanically mixed isothermal bath.

The 0.37 Hz peak is seen as a "grouping" near the "expected" frequency that correlates with the measured side-branch spacing and tip growth rate. This particular peak (or group of peaks) becomes more significant as data farther behind the tip (i.e., closer to the side branches) is included in the measurements. This strengthening of the signal with sampling region size reinforces the interpretation of this peak as being related to the side-branching process, even thought the data derives from a region much closer to the tip than the side-branch structure proper. This constitutes evidence that the tip itself is not stationary in the "tip frame of reference", and is instead operating in a limit cycle mode. Earlier work by Dougherty *et al.* [13], suggested that such a limit cycle was not present in NH_4Br dendrites, nor did side branches on opposite sides of the dendrite appear to correlate (spatially) with each other. These issues will be examined in more detail as this investigation continues.

5 Conclusions

In summary, a method was developed for evaluating dendritic growth rates that discriminates fine distinctions in the non-steady-state behavior. This method is applied here to PVA dendrites, seen to grow in microgravity at non steady-state velocities. The mechanism responsible for this behavior may be related

to thermal interactions between a dendrite and its surroundings, or it may be intrinsic to the dendritic solidification process. Efforts to further discriminate among the possible causes are currently under way.

Spectral analysis of the dendritic growth data reveals a characteristic frequency that may be fundamental to the growth process. A peak is seen in the presented growth near the frequency at which side arms are being created (0.37 Hz), the significance of which increases when data further behind the tip is used in the analysis. Nonetheless, this peak is present even when the analyzed data is derived from the smooth, steady-state region of the interface, 10-times closer to the tip than the first measurable side-branch (via microscope).

The steady-state aspects of the dendritic growth process are reasonably well understood. However, the range of non-steady-state behaviors observed in these experiments and others, illustrates that much further examination is needed. Towards this end, we will continue our efforts to identify and better understand the time-dependent aspects and how they contribute to the transport of heat and mass, the physics of interfacial stability, and the more general behavior of non-isolated dendrites in engineering systems.

Acknowledgements

This work was supported under NASA Contract Nos. NAS3-25368 and NAG8-1488. The authors wish to thank D.P. Corrigan, L.A. Tennenhouse, and the many Rensselaer undergraduate and graduate student volunteers for assistance in operating the flight experiment. Additional thanks to IDGE Project Manager D.C. Malarik and her team at (or associated with) NASA's Glenn Research Center at Lewis Field for their engineering support, and the staff of the POCC/HOSC at NASA's Marshall Space Flight Center for flight support. Finally we thank the crew of the STS-87 for bringing our experiment to orbit and back, and for their extra effort in providing us with the additional best available quality video data for our post flight analysis.

References

1. M.E. Glicksman and S.P. Marsh: 'The Dendrite'. In: *Handbook of Crystal Growth.* ed. by D.T.J. Hurle (Elsevier Science Publishers, Amsterdam 1993) pp. 1077
2. U. Bisang and J.H. Bilgram: Physical Review E, **54**(5), 5309 (1996)
3. G.P. Ivantsov: Dokladi Akademii Nauk, USSR, **58**(4), 567-569 (1947)
4. S.-C. Huang and M.E. Glicksman: Acta Metallurgica, **29**, 701-715 (1981)
5. V. Pines, A. Chait, and M. Zlatkowski: Journal of Crystal Growth, **167**, 383-386 (1996)
6. V. Pines, A. Chait, and M. Zlatkowski: Journal of Crystal Growth, **182**, 219-226 (1997)
7. J.C. LaCombe, M.B. Koss, D.C. Corrigan, A.O. Lupulescu, L.A. Tennenhouse, and M.E. Glicksman: Journal of Crystal Growth, **206**(4), 331-344 (1999)

8. J.C. LaCombe, M.B. Koss, D.C. Corrigan, A.O. Lupulescu, J.E. Frei, and M.E. Glicksman: 'The Influence of Interface Morphology on Heat Transport in Dendritic Solidification'. In: *Solidification 1999*. ed. by W.H. Hofmeister, *et al.* (TMS, Warrendale, PA 1999) pp. 121-130
9. N. Provatas, N. Goldenfeld, J. Dantzig, J.C. LaCombe, A. Lupulescu, M.B. Koss, et al.: Physical Review Letters, **82**(22), 4496-4499 (1999)
10. M.B. Koss, J.C. LaCombe, L.A. Tennenhouse, M.E. Glicksman, and E.A. Winsa: Metallurgical and Materials Transactions A, **30A**, 3177 (1999)
11. 11.N.R. Lomb: Astrophysics and Space Science, **39**, 447 (1976)
12. W.H. Press, S.A. Teukolsky, W.T. Vellerling, and B.P. Flannery: 'Numerical Recipes in Fortran' , 2nd edn. (Campridge University Press, New Delhi 1994) pp.
13. A. Dougherty, P.D. Kaplan, and J.P. Gollub: Physical Review Letters, **58**, 1652-1655 (1987)
14. J.C. LaCombe, M.B. Koss, and M.E. Glicksman: Physical Review Letters,**83**(15), 2997-3000 (1999)
15. J.C. LaCombe, M.B. Koss, V.E. Fradkov, and M.E. Glicksman: Physical Review E, **52**(3), 2778 (1995)

Patterns in the Bulk and at the Interface of Liquid Crystals

Ágnes Buka[1], Tamás Börzsönyi[1,2], Nándor Éber[1], and Tibor Tóth-Katona[1]

[1] Research Institute for Solid State Physics and Optics, Hungarian Academy of Sciences, H-1525 Budapest, P.O.Box 49, Hungary
[2] Groupe de Physique des Solides, CNRS UMR 75-88, Universités Paris VI et VII, Tour 23, 2 place Jussieu, 75251 Paris Cedex 05, France

1 Introduction

Liquid crystals (LCs) are substances possessing one or more mesophases between their liquid and solid phase. The sequence of these mesophases represents a step by step ordering of the structure. The nematic (N) phase is characterized by an orientational order of the elongated molecules described by the director $\mathbf{n}$, but the centers of mass of the molecules are arranged randomly. In the smectic (Sm) phases besides the orientational order the centers of mass of molecules form a layered structure. The smectic-A (SmA) phase has no positional order within the layers, while the smectic-B (SmB) phase is characterized by a long-range hexagonal order in each layer and by a weak correlation between the layers. The features of these and other LC phases are described in detail in the literature – see e.g. [1,2,3].

Liquid crystals are quite rich in pattern forming phenomena [4]. Owing to the orientational order LCs possess strongly anisotropic physical properties, which on the one hand result in the known pattern forming processes (observed in other systems such as Rayleigh-Bénard convection, viscous fingering, etc.) with richer scenarios, and on the other hand allow for new instability mechanisms (such as electrohydrodynamic convection, patterns at the Freedericksz transition, etc.). Furthermore, optical and other properties of LCs (e.g. birefringence, transparency, phase transition temperatures close to the room temperature, relatively fast transport processes) make the detection of patterns and realization of the experiment easier. Therefore LCs are often considered as model materials for studying both bulk and interfacial instabilities. In general, patterns in LCs are characterized by a relatively low threshold, high regularity and by a large aspect ratio (LC sandwich cells typically have lateral dimensions of $x, y \approx 1-2$ cm and thickness of $z \approx 10-100$ μm).

2 Bulk Instabilities

When a liquid crystal is confined between properly treated surfaces usually a homogeneous equilibrium (basic) state is obtained. In this state the elastic restoring torques force all spatially periodic director fluctuations to decay. External influences (electric or magnetic fields, thermal gradients or mechanical stresses) may,

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 298–318, 2001.
© Springer-Verlag Berlin Heidelberg 2001

however, lead to the appearance of destabilizing bulk torques. Then if the external field exceeds some critical threshold value the destabilizing torques may overcome the restoring ones, i.e. the amplitude of fluctuations starts growing and a nonhomogeneous pattern evolves. During this process typically a wavelength selection occurs, thus a pattern (stripes, spirals, targets, hexagons) is obtained which is characterized by a spatially periodic modulation of the director. The anisotropic optical properties of liquid crystals make these patterns easily observable either in microscopes with or without polarizers or via laser diffraction.

The above considerations are almost independent of the actual mechanism of the instability hence various physical processes, as e.g. the thermally driven Rayleigh-Bénard instability, the electric field induced convection or the shear flow instabilities, may provide very similar patterns [4].

2.1 Electroconvection

Electroconvection (electrohydrodynamic instability) [5] is a pattern forming process which – as explained by Carr and Helfrich in the late 60s – is related to the anisotropic physical properties of the liquid crystal.

In the classical electroconvection (EC) experiments a *planarly oriented* ($\mathbf{n} \| x$) layer of a nematic liquid crystal with negative dielectric anisotropy ($\varepsilon_a < 0$) corresponds to the homogeneous basic state. Besides the elastic torques due to the director gradient in this geometry the dielectric torques (which tend to keep the director perpendicular to the electric field) also serve as restoring torques. Due to the anisotropy of the electrical conductivity ($\sigma_a > 0$), however, in the presence of a director modulation (a thermal fluctuation) the electric currect may have a component perpendicular to the field which leads to charge separation. The electrostatic force acting on the charges induces a convective flow which provides a viscous destabilizing torque on the director.

Patterns at the Onset of the Electroconvection

The EC instability has been extensively studied since the first observation of a regular roll pattern, the Williams domains [6]. The formation of patterns is governed by two control parameters, the amplitude V and the frequency f of the applied voltage. Varying the frequency two regimes can be distinguished [5]. At low frequencies (in the 'conductive' regime) the periodic director and flow patterns are stationary and the charges oscillate with the field. At high frequencies (in the 'dielectric' regime) on the contrary, the charge distribution is stationary while the director and the flow velocities oscillate. The two regimes are separated by the cut-off frequency f_c which is roughly proportional to the electrical conductivity. In the 'conductive' regime there can be two different scenarios. The direction of the rolls is either perpendicular to the director (normal rolls, NR), or it is tilted (oblique rolls, OR) forming zig and zag domains. Oblique rolls can be observed at frequencies below the Lifshitz point f_L, with the roll angles increasing with decreasing frequency.

The threshold voltage V_c and the critical wavevector $\boldsymbol{q_c}$ of the roll pattern at the onset of the instability can be calculated by a linear stability analysis of the standard model of EC (a set of 6 coupled differential equations for the director, the flow velocity and the charge distribution) [7,5]. This model provides a full description of all above scenarios in agreement with the experiments.

In the 'conductive' regime V_c has a strong frequency dependence (it diverges approaching the cut-off frequency), while the wavelength λ_c of the rolls is about the sample thickness d. In the 'dielectric' regime the threshold grows with the square root of f, while the wavelength - only a few μm - is independent of the thickness and for typical cells ($d > 10$ μm) is much smaller than in the 'conductive' regime. In both regimes λ_c decreases with increasing frequency. One usually introduces $\epsilon = (V^2 - V_c^2)/V_c^2$ as a reduced control parameter characterising the deviation from the threshold.

Roll patterns can be observed in *twisted planar* cells too, where at the two boundaries the director is oriented perpendicular to one another. In this case the rolls are expected to be directed at 45 degrees (along the bisectrix of the preferred directions of the surfaces). This has been proved for the 'conductive' regime, however, according to recent experiments it might not always hold for 'dielectric' one. Diffraction measurements have shown a superposition of two roll structures above some critical frequency which depends on the sample thickness [8]. These rolls are oriented along the preferred orientations of the surfaces and are thus normal to one another. This observation indicates a possible transition from the standard bulk electroconvection state (rolls at 45^o) to an instability which is restricted to thin regions at the surfaces. Though such surface roll structures have been proposed earlier as a model for the 'dielectric' rolls [9], no theoretical explanation has yet been given for this transition which seems to occur at a small wavelength to thickness ratio.

In *homeotropic* cells ($\mathbf{n} \parallel z$) in electric field the first instability is a homogeneous tilt of the director (a Freedericksz transition – FT [1,3]), however, EC may set in at higher voltages. As the homeotropic orientation is fully degenerate in the surface plane, the direction of the tilt and hence that of the rolls are randomly selected and may vary in space and time. That represents a direct transition to spatio-temporal chaos already at the onset of the EC (soft mode turbulence) and is therefore subjected to detailed experimental and theoretical investigations [10,11,12,13,14].

The azimuthal degeneracy in the xy plane can be removed by an additional magnetic field $\boldsymbol{H}$ parallel to the surfaces (i.e. perpendicular to the electric field) which selects the tilt direction [11,13,15]. Under such conditions the EC scenarios typically are very similar to those in the planar case. It has been shown, however, both experimentally and theoretically that certain combination of the material parameters may lead to the appearance of a second (low frequency) Lifshitz point [16,17], i.e. to the preference of normal rolls at very low frequencies. For the time being this feature has only been observed in homeotropically oriented samples of Nematic Phase 5A (Merck).

Scenarios above Threshold

Until now we focused on the characteristics of the EC patterns at the onset of the instability. Increasing the voltage above V_c will, however, modify some of these characteristics (vary the wavevector or the director profile) and may also lead to new scenarios.

One of the scenarios at higher ϵ is the *'abnormal' roll* (AR) pattern which develops from normal rolls. As shown experimentally first in homeotropic [18,19], [15,16,17], later also in planar cells [20,21], in 'abnormal' rolls the director may rotate out of the tilt plane gaining a $\varphi \neq 0$ azimuthal angle, while the direction of the rolls remains unchanged. In homeotropic cells the appearance of such an out-of-plane component of the director leads to a net rotation of the optical axis of the sample, thus the 'abnormal' rolls can be easily detected using crossed polarizers (Fig. 1). In planar cells, however, due to the strong surface anchoring only a twist deformation can develop, thus detection requires more sophisticated techniques [22,21].

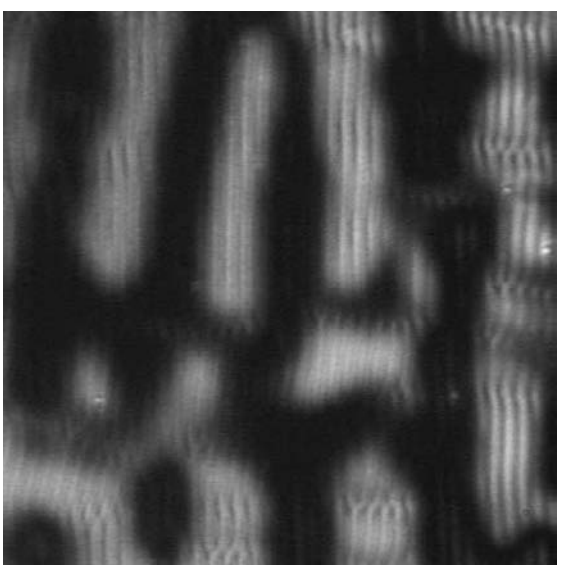

Fig. 1. 'Abnormal' roll domains (dark and bright patches) in a homeotropic nematic Phase 5A (Merck) in the presence of a small orienting magnetic field. Crossed polars are rotated by 8^o with respect to the magnetic field.

The equations of the standard model of EC are too complicated to be solved in the nonlinear regime. Instead, for a weakly nonlinear analysis a Ginzburg-Landau formalism is used, i.e. an amplitude equation is constructed [23,5]. Inspired by the experiments a reconsideration of the symmetries of the system has led to the realization of the fact that the pattern amplitude in itself is not enough to describe the nonlinear behaviour. Allowing for a Goldstone mode one has to

take into account an additional degree of freedom, the azimuthal angle φ of the director, which is coupled to the pattern amplitude A [24,20]. This feature is a consequence of the anisotropy of liquid crystals so it is not a privilege of EC, the coupling between A and φ is observable in thermal convection too [25]. Solving the resulting coupled differential equations thresholds of secondary instabilities could be derived and an $\epsilon - f$ phase diagram could be constructed. It showed that the NR pattern becomes unstable either due to a zig-zag instability (modulation of the roll direction) near the Lifshitz point or by a normal–'abnormal' roll transition at higher frequencies. In this latter case the out-of-plane component of the director appears as a forward (pitchfork) bifurcation for the azimuthal angle φ with increasing ϵ. As the rotations into both out-of-plane directions are equally probable, two types of domains with opposite sign of φ are expected to exist. The phase diagrams in the planar and the homeotropic (with a small superposed orienting magnetic field) geometry look similar except that in homeotropic samples the critical ϵ for the secondary instabilities are much smaller and the frequency range for the zig-zag instability is narrower. The theoretical predictions could be checked experimentally for both geometries and an excellent agreement has been found [21,16]. However, in contrast to the expectations, in homeotropic samples no distinct domains with sharp boundaries could be detected, instead a continuous (almost periodic) variation of φ could be observed with an approximate wavelength about 10 times that of the rolls [16,17] – 'prechevron' state (see later).

Increasing the voltage much above V_c the regularity of the roll pattern decreases by the appearance of *defects* which are typically dislocations in the convection roll pattern. The number of defects increases with the voltage as well as their mobility until we enter a chaotic, turbulent regime.

These defects play an important role as any adjustment of the wavevector can be realized only by generation and motion of defects. Changing the control parameters often leads to conditions where the actual wavevector $\mathbf{q}$ of the pattern is substantially different from the ideal $\mathbf{q_{id}}$ one belonging to the given voltage and frequency. If the wavevector mismatch $\mathbf{\Delta q} = \mathbf{q} - \mathbf{q_{id}}$ is large enough defects are usually generated in pairs and start moving to reduce $\mathbf{\Delta q}$. Motion of a defect along the roll direction ('climb') will modify the roll spacing, while motion normal to the rolls ('glide') can adjust their direction. Using the Ginzburg-Landau formalism the velocity $\mathbf{v}$ of an individual (far from and therefore not interacting with their neighbours) defect could be calculated yielding $\mathbf{v} \perp \mathbf{\Delta q}$ and a logarithmic singularity in v for $\Delta q \to 0$ [26]. In early experiments on planar EC only the 'climb' motion could be excited by sudden simultaneous change of the voltage and frequency [27,28]. In the homeotropic geometry the 'glide' motion can easily be generated by rotating the small orienting magnetic field. In recent experiments using local heating due to laser pulses for generating the defects both the predicted direction of $\mathbf{v}$ and the logarithmic singularity of $v(\Delta q)$ could be verified [17,29].

It has been known for a long time that in the high frequency EC (in the 'dielectric' regime) the regular roll pattern breaks down at higher voltages to

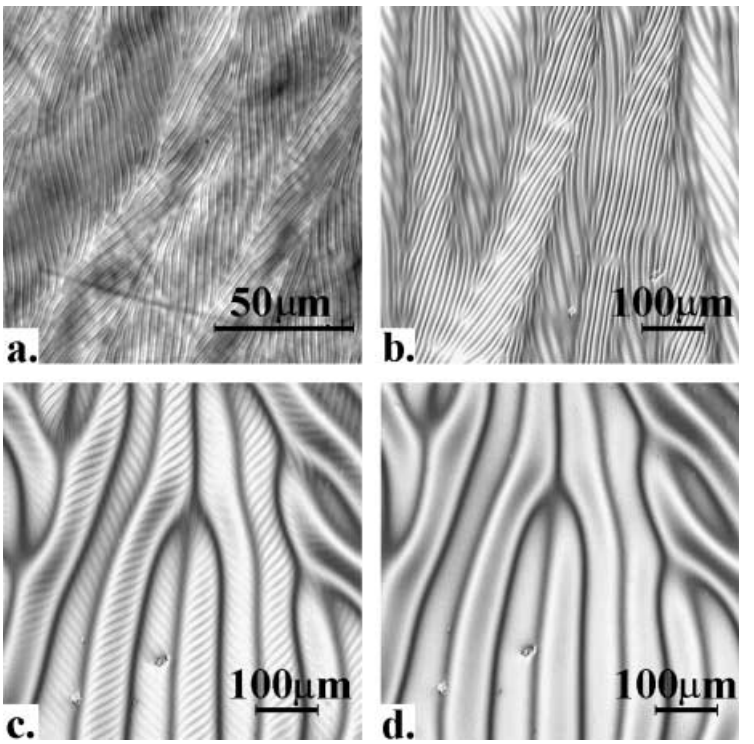

Fig. 2. Chevron scenarios in electroconvection of the homeotropic nematic Phase 5A (Merck) . a. 'dielectric' chevrons, b. 'conductive' defect mediated chevrons, c. 'conductive' defect free chevrons, d. prewavy pattern. Crossed polars. Sample thickness: a. $d = 20$ μm, b-d. $d = 60$ μm.

chevron scenario [30,31], i.e. to a periodic arrangement of domains. In the neighbouring domains the dielectric rolls are rotated alternatingly with respect to their initial direction (zig and zag chevron domains, Fig. 2a). It has been explored that the domain boundaries in these dielectric chevrons consist of an ordered chain of defects (dislocations) running along the initial roll direction [28,32,33], moreover, besides the rolls the orientation of the director is rotated too [22]. Thus in the chevron pattern the azimuthal angle φ of the director is periodically modulated which can be detected optically just as in the case of the 'abnormal' rolls in the 'conductive' regime. This φ modulation remains observable even if the visibility of the individual 'dielectric' rolls is much reduced either due to their small wavelength or to the bulk-surface transition mentioned earlier.

Recently the ordering of defects into chains has been theoretically explained [34]. It has been shown that its mechanism is not restricted to the 'dielectric' regime, under proper conditions similar chevrons can exist in the 'conductive' regime too. Such 'conductive' chevrons have actually been observed

in homeotropic cells, especially in the absence or with only a small orienting magnetic field (Fig. 2b) [15,17]. This scenario, which has been classified as a *defect mediated chevron* [35,36] reminding the role of defects, typically occurs at voltages much above the NR-AR transition.

Both the 'dielectric' and the 'conductive' chevrons possess a dual spatial periodicity, one for the convective ('dielectric' or 'conductive') rolls, the other for the defect chains. The wavelength of the latter is typically about an order of magnitude larger than that of the rolls, i.e. it can be larger than d.

These chevrons are not the only patterns with a wavelength larger than d. Such patterns have been reported for MBBA strongly doped with ionic dopants in planar and twisted planar ('wide' domains) [37,38,3], as well as in homeotropic geometries (*'prewavy' pattern*) [31,32,35,36]. These large wavelength patterns have in common that they evolve from a (seemingly) pattern free state at a threshold voltage having a weaker frequency dependence compared to either that of the 'conductive' or the 'dielectric' rolls. They can be observed with crossed polars (Fig. 2d), but the shadowgraph image with a single polarizer is almost invisible, in contrast to the usual convection rolls. That indicates an only azimuthal (without tilt) modulation of the director. It has been measured recently in homeotropic cells that the amplitude of the azimuthal modulation in the prewavy pattern increases with the voltage resembling a supercritical pitchfork bifurcation [36,39]. At higher voltages a transition to chevrons could be observed. However, while at higher frequencies these did correspond to the 'dielectric' chevrons mentioned above, near the Lifshitz point a different type of 'conductive' chevron pattern was observed (*defect free chevrons*) [35,36]. Though it looks similar to defect mediated chevrons, no defects (dislocations) are present, instead a continuos curvature of the rolls is observed (Fig. 2c). The chevron domains coincide with the stripes of the initial prewavy pattern. These defect free chevrons can thus be interpreted as 'conductive' normal rolls with azimuthally modulated director orientation.

The appearance of the 'wide domains' or the prewavy pattern has not been theoretically understood yet. At present no physical mechanism is known which could produce such a periodic static deformation (without flow) at thresholds lower than that of the electroconvection. The observed motion of dust particles in these patterns (and even at voltages below threshold) stress the importance of convection [40]. It suggests that an electroconvection mechanism, though different from the Carr-Helfrich one, might be responsible for the formation of these patterns [41]. One possibility could be the 'isotropic' mode [42] which is expected to produce very fine rolls unresolvable by the visible light, however, the chevrons formed from the defects of these fine rolls would possess similar features as the prewavy pattern. Unfortunately, no direct experimental evidence has been found to prove this assumption yet.

The defects (dislocations) mentioned so far usually possess a well defined core. Sometimes, however, one can observe *'long' defects*, where the core is delocalized to phase–jump lines extending over several rolls (Fig. 3a). Computer simulations

have shown that such structures are quite common in the AR range both in the planar and in the homeotropic geometry [43].

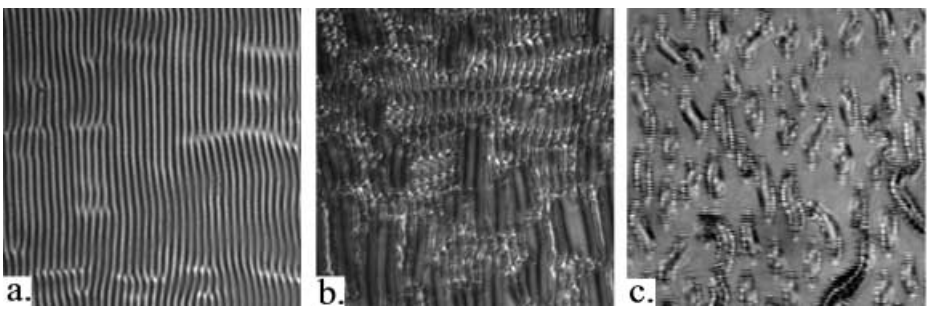

Fig. 3. Exotic patterns in electroconvection of nematics. a. 'long' defects and b. coexisting abnormal and CRAZY rolls in homeotropic Phase 5A (Merck), c. localized structures (worms) in planar 10E4.

Besides the defects in the convection rolls (dislocations, 'long' defects) the pattern may contain orientational singularities of the director field which usually appear in the form of domain walls or *disclination loops*. These orientational defects are quite common in homeotropic nematics. One recently reported example is the CRAZY roll pattern (Fig. 3b) [16,17] which may appear near the Lifshitz point above the normal-abnormal roll transition. The CRAZY rolls correspond to closed disclination loops located in the plane perpendicular to the cell surfaces and parallel to the rolls, which grow into the abnormal roll pattern without changing the wavelength or causing dislocations. Disclination loops play an important role in the Zvinger pattern [44] observed in a substance possessing a nematic-smectic phase transition as well as during the prewavy-wavy pattern transition. All these patterns appear high above threshold far beyond the applicability of either the linear or the weakly nonlinear stability analysis.

Another fascinating phenomenon in the nonlinear range is the spontaneous formation of localized states or pulses ('worms'). Chaotic localized traveling-wave states which coexist with the 'conduction' state, are long lived, appear and die at irregular locations and times, have unique small widths and irregularly varying length, as seen in I52 [45]. Time dependent localized states ('roll sinks' and 'butterflies') coexisting with the uniform state have been observed in 10E6 [46] (see an example on 10E4 in Fig. 3c).

2.2 Shear

Shear flow phenomena in nematic liquid crystals have drawn much attention to the experimental side, since the orientational instabilities are easy to be observed, and they take place at much smaller shear rates than the laminar-turbulent transition. This results from the fact, that the relaxation of the director fluctuations is usually orders of magnitude slower than that of the velocity fluctuations.

The basic feature of the instabilities is, that the flow (velocity) field $\mathbf{v}(r,t)$ and the director $\mathbf{n}(r,t)$ is coupled in the governing nemato-hydrodynamic equations. For most of the substances due to the orienting effect of the flow in the stationary case the director encloses a small angle Θ_{fl} with $\mathbf{v}$ (flow alignment) [1]. For some nematic liquid crystals (in particular near a nematic-smectic transition) no equilibrium state exists, but the flow exerts a continuous torque on the director, which might lead to a tumbling motion [47]. Here we will consider the case of flow aligning materials.

Rectilinear Flow

The instability mechanism varies considerably with the type of shear (stationary/oscillatory and Couette/Poiseuille flow), and each case has been considered both from the theoretical point of view and from the experimental side with a large variety of results which are summarized in [48] up to 1995.

The two basic cases are when the director is aligned (with external fields, or surface treatment) perpendicular to the shear plane (defined by $\mathbf{v}$ and $\nabla\mathbf{v}$) or in the shear plane. In the first case there is no torque on the director, but the configuration is unstable and above a threshold perturbations grow. The mechanism has been described by Pieranski and Guyon and was investigated in details [49]. In the second case the flow field exerts a torque on the director, and its orientation results in a balance between the elastic (or electric/magnetic) torques and the torque applied by the flow.

From now on we consider only the case of oscillatory flow fields (with the frequency f) that can be more easily realized experimentally than the case of stationary flow. One may first consider the response of the system before any instability occurs (basic state) where the director generally oscillates homogeneously around its initial equilibrium position. Beyond the small-amplitude linear regime one finds a response that is nonlinear in the shear strain. The temporal behavior of the director has been calculated using some approximations [50]. In the case of Couette flow (the shear is applied by moving one of the bounding plates in its plane) at low frequencies (f below about 400 Hz) the viscous penetration depth $\sqrt{\eta/\rho\omega}$ ($\omega = 2\pi f$), is much larger than the cell thickness d, where η is an appropriate effective shear viscosity and ρ denotes the mass density of the nematic. Then the flow can be approximated by the simple linear flow field, which amounts to neglecting the time derivative (inertia term) in the Navier-Stokes equation. With this assumption one may expect Freedericksz-type instabilities where the time-averaged director reorients homogeneously in the plane of the layer. For the simple linear Couette flow field no spatially homogeneous instabilities are predicted to occur (in contrast to the case of Poiseuille flow) [51,52], even if the possibility of transitions out of the flow plane is included.

Above a critical flow amplitude A_{xc}, transitions to spatially periodic roll states have been observed in homeotropic (director perpendicular to the confining plates) [53,54] and planar alignment [55]. In Fig. 4.a a developed roll pattern is shown in the homeotropic alignment [54].

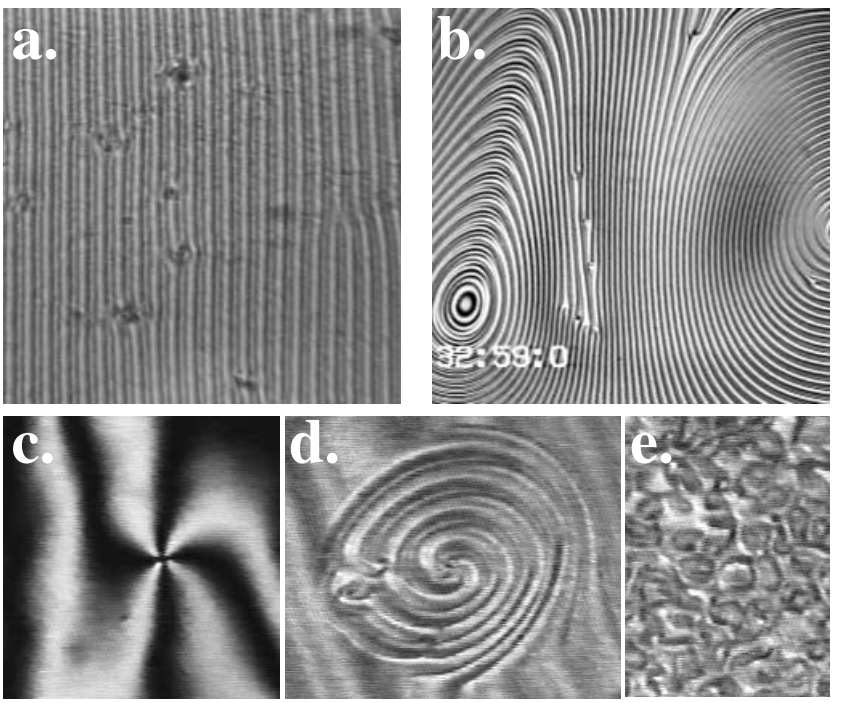

Fig. 4. **(a.)** *Rectilinear shear:* Roll pattern observed in polarized white light at $f = 80$ Hz and $A_x/d = 0.5$ in 5CB. The rolls are perpendicular to the direction of the upper plate oscillations. **(b.)** *Oscillatory compression of the sample:* A typical "target like" pattern in Phase 5 (Merck) at f=75 kHz for $d = 20$ μm. **(c-e.)** *Elliptic shear:* The spatio-temporal behavior of the system is considerably different in the three regimes: **(c.)** At large electric fields a nearly homogeneous precession is observed; **(d.)** Around $U = 1.2U_F$ spiral formation has been detected; **(e.)** At smaller voltages ($U < 1.2U_F$) spatiotemporal chaos can be observed. Phase 5 (Merck), f=155 Hz, $A_x = A_y = 3.4$ μm, and $d = 20$ μm.

The threshold amplitude of the roll pattern decreases with f [54]. A sufficiently rigorous linear stability analysis (that goes beyond the lowest-order modes) gives a good quantitative agreement with the experimental threshold values except for parameter ranges where the critical oscillation amplitude (in physical units) becomes large [54].

When comparing our data with the roll threshold obtained in the planar geometry [55] we find that the roll instability develops at smaller amplitudes in the homeotropic geometry than in the planar, as it is expected intuitively (the planar geometry is closer to the preferable flow alignment). However, the value of A_{xc} differs by a factor of 2, its frequency dependence is similar.

Director Precession, Phase Waves

An other interesting case is, when an elliptic shear is applied by the linear oscillation of the bounding plates in their plane perpendicular to each other. At large enough amplitudes and small ellipticities similar types of orientational instabilities take place as for the case of rectilinear flow. By increasing the ellipticity (towards the circular case) a transition from the roll regime to hexagonal patterns is observed [57].

An interesting configuration is, when the director is already tilted out by an external electric field from the initial homeotropic orientation ($\varepsilon_a < 0$), and a small amplitude elliptic shear is applied (below the instability threshold) [58,59]. The director is forced to precess around the z axis by the shear with a frequency Ω which is normally orders of magnitude smaller than f.

The value of Ω increases linearly with f and $A_x * A_y$ [59]. This observation is in accordance with the analytical calculations – which are done for low frequencies ($f < 10^4$ Hz), small amplitudes ($A/d \ll 1$), circular shear, and neglecting spatial dependence in the xy plane – yielding $\Omega \approx (\lambda/2)(A/d)^2\ 2\pi f$ where $\lambda = \alpha_3/\alpha_2 \approx 0.05$ for usual nematics.

The dependence of Ω on U is more complicated and for its understanding probably one has to take into account the spatial dependence. At sufficiently high voltages the director orientation varies slowly in space and precesses almost homogeneously in time (Fig.4.c). Around $1.2U_F$ inhomogeneities emit traveling waves and umbilics generate spiral waves (Fig.4.d), very similar to those observed in oscillatory and excitable chemical reactions [60]. At lower voltages one observes spatio-temporal chaos (Fig.4.e).

We have also performed experiments with linear mechanical vibration along z (compression) generated by piezo crystals ($f \approx 5$ - 100 kHz) attached to one of the bounding plates. The z oscillations presumably induce Poiseuille flow. As before, the slow precession occurred only in the Freedericksz distorted state.

Here the phase waves are typically emitted from certain locations in the form of target patterns (see Fig. 4.b), which presumably result from spatial inhomogeneities in the flow. The waves behave diffusively (even near U_F), which is probably due to the fact that at the high frequencies used the elastic contributions to the precession are irrelevant. In the immediate neighborhood of the FT and at not too high temperatures, a reversal of the precession was observed in Phase 5 and MBBA. For more details, see [61].

Director precession and phase waves have been observed previously in cells that were excited piezoelectrically at frequencies around 50 kHz [62]. The piezo crystal formed one of the bounding plates. Phenomena reminiscent of the phase waves were also seen in planar and homeotropic cells without electric field at frequencies 10 kHz $< f <$ 1 MHz [63]. There the waves originated from orientational defects at the surface.

3 Interfacial Instabilities

Interfacial instabilities appear at the interface that divides two different substances, phases or structures. The interface is driven by the gradient of a field which itself obeys a diffusion or Laplace equation. The emergence of the pattern results from the interplay between the stabilizing effect of the surface tension and interface kinetics and the destabilizing effect of the diffusive field.

From mathematical point of view these systems are similar. The surface tension causes curvature effects that should be incorporated into the equation describing a local equilibrium at the interface (moving boundary condition). Consequently, it is not surprising that all these systems show very similar patterns and that in the instability mechanisms the same parameters play important roles, i.e. the anisotropy of the surface tension and the interface kinetics.

Interfacial instabilities have been investigated extensively theoretically as well as experimentally in the last decades [64,65]. On the experimental side solidification [66], viscous fingering [67], electrochemical deposition [68], as well as growth of bacterial colonies [69] have been investigated. On the theoretical side mostly numerical approaches are used, since the analytical treatment of the problem is rather difficult except the stability analysis of the few (planar, circular or dendritic) solutions [65]. The two basic numerical methods are the sharp interface models (direct tracking of the front) [70,71] and the phase field models which are dynamic extensions of the Cahn-Hilliard theory of first-order phase transformations [72].

3.1 Propagation of a First Order Phase Transition Front

Here the growth is limited by diffusion (of heat or impurities) that exerts a destabilizing force on the (macroscopically) smooth interface. It is the surface tension and the phase transformation kinetics which tend to stabilize the interface against perturbations. Several liquid crystalline phase transitions are suitable to study the interfacial instability mechanisms:

The nucleation of the SmA phase in the isotropic (I) phase leads to the growth of battonets (rod like shape) and if they are bent the focal conic textures [2] develop. The formation of battonets is understood by taking into account the free energy difference between the two phases, the elastic energy and the surface energy [73].

The other well investigated phase transition is the $I \rightarrow$ columnar hexagonal transition [74,75]. In the columnar phase rod like objects (columns) are formed by the packing of disc like molecules on top of each other. A hexagonal order is present between the columns. The morphology diagram was reported and experimental results on both the surface tension and the kinetic anisotropy have been given [74].

Dendritic growth has been observed in the $I \rightarrow N$ [76] and $SmA \rightarrow SmB$ [77] phase transitions only in directional solidification.

In the smectic to crystal phase transition the morphological changes and the nature of mode selection were analyzed in [78].

Detailed studies have been made on the $N \rightarrow SmB$ transition [79,80,81,82,83] on a homologous series CCHm (m=3; 4; and 5) of a bicyclohexane compound. A thin (quasi 2D) sample after equilibration was cooled down below the phase transition temperature T_{NS} and the form of the nucleating and growing SmB domains was detected and analyzed as a function of the undercooling $\Delta T = T_{NS} - T$. Other details of the experimental setup can be found in [79]. With a proper experimental procedure both planar and homeotropic orientations have been achieved in the SmB and N phases and by that different anisotropies have been assured in the plane of observation.

Since the growth was observed in a quasi two-dimensional geometry, the angular dependence of the interfacial properties are relevant only in the plane of the sample.

As it can be expected, the anisotropies in a plane perpendicular to the smectic layers are much larger than parallel to them.

It has been shown [79,81], that in the case of a planar smectic germ **(A.)** the anisotropies resulted from the structure of the smectic phase are dominant in the pattern formation, while in the case of a homeotropic smectic germ **(B.)** the effects originating from the anisotropy of the nematic phase become comparable to the effects originating from the anisotropy of the smectic phase.

(A.) When the smectic layers are perpendicular to the sample (planar smectic germ) at low undercooling the formation of facets is observed that are parallel to the layers.

Normally the difference between the dynamics of a system with small anisotropies and a system showing facets can be summarized as follows. In a system with small anisotropies the solidification front is rough on the atomic level and the phase transformation (attachment of particles to the crystal) is limited by heat and impurity diffusion [65]. On the other hand faceted interfaces are smooth on an atomic scale and the growth of the solid phase is limited by the attachment kinetic process [84]. Thus in the later case usually the facet is blocked (does not advance) even at large undercoolings.

According to our observations this is not the case for the facets at the $N - SmB$ interface of CCH3, where even at very small undercoolings ($\Delta T = 0.02$ ^{o}C) the facet advances with a finite velocity. (The case of CCH4 and CCH5 is more complicated, here the propagation of the facet was only detectable at $\Delta T > 0.1$ ^{o}C which is still a small undercooling.)

Preliminary directional solidification measurements (a quasi-two-dimensional sample is pulled with a constant speed in a constant thermal gradient towards the cold side) also show that the nematic - smectic B front is more stable and variations in the facet velocity are much smaller than for traditional faceted fronts (e.g. the nematic - crystal interface of the same material) [85].

Phase-field simulations have shown [80], that even in the extreme (probably not realistic) case (i.e. supposing isotropic interfacial kinetics) with a "cusp-like" angular dependent surface tension obtained by the Wulff analysis of the nearly equilibrium shapes of the smectic-B domains, at small undercoolings the growth shapes are faceted (at a macroscopic scale) and at larger undercoolings

four-armed dendrites are formed, which is in accordance with our experimental observations [79].

(B.) To study the effect of the anisotropy of the nematic phase (liquid side) in the interfacial instabilities the most suitable configuration is when the smectic layers are parallel to the quasi two-dimensional layer of investigation, since the anisotropy of the SmB phase is the smallest in the plane of the smectic layers.

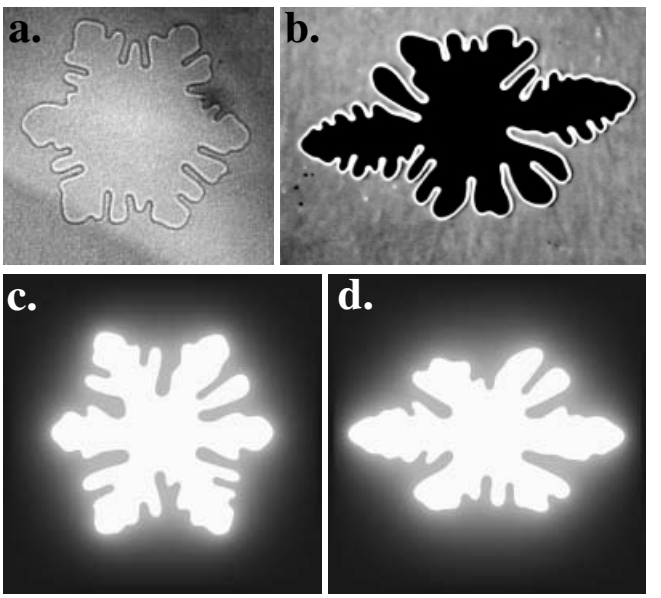

Fig. 5. The growth shapes of homeotropic smectic domains in homeotropic (**a.**), and in planar (**b.**) nematic phase at $\Delta T = 0.2$ °C and $\Delta T = 0.15$ °C respectively; Simulated growth shapes (using the phase field model) in isotropic (**c.**) and anisotropic (**d.**) liquid phase.

The growth morphologies have been compared for homeotropic and planar nematics [81] (see Fig. 5.a-b). Using the phase field model we have analyzed separately the effects coming from the superposing anisotropies of the surface tension, phase transformation kinetics and heat diffusion in the nematic phase.

The most interesting result is, that the heat diffusion anisotropy induces an elongation of the germ and the formation of dendritic tips perpendicular to the nematic director $\mathbf{n}$(N) [81] (See Fig. 5.a-b.).

This effect has also been observed in the case of planar smectic domains in CCH4 where the two opposite pairs of the main arms of the four-armed dendrites grow with slightly different velocities (asymmetric growth) if $\mathbf{n}$(N) is not a symmetry axis of the growth directions [82].

Finally, one should mention that material parameters of CCH3 (T_{NS}, volume change on phase transition, latent heat, electric conductivity, etc.) are in the

range that allows an easy control of the dendritic side-branching process by pressure oscillations or by a periodic heat release in the volume [83]. Such a method of controlling the side-branching process besides of academic aspects might have also implications in practical applications.

3.2 Viscous Fingering

The Saffman-Taylor (viscous fingering) problem, where a less viscous liquid (or gas) is pressed into a liquid with larger viscosity, represents a relatively simple case of nonlinear interfacial pattern formation and has been intensively studied – see e.g. reviews [86,67]. The evolution of the air – viscous fluid interface placed between two close parallel plates (Hele-Shaw cell) is determined by the pressure field p in the fluids satisfying the Laplace equation subjected to boundary conditions that are the pressures far from the interface on its both sides, and the pressure drop across the interface given by the Gibbs-Thomson condition and by the wetting properties.

For the interface of isotropic fluids experimental results agree well with the linear stability analysis [87,88]. Later stages of the interface growth in the radial geometry (after breakup of the initial circle) are less understood than those of the channel flow. For isotropic fluids the pattern continues to evolve through repeated tip-splitting to form more and more fingers.

Using non-Newtonian fluids (e.g. liquid crystal) as (usually) a more viscous fluid [89,90] however, gives a much richer morphological diagram than that obtained for isotropic fluids, due to the *inherent anisotropy*. As mentioned previously, in nematic liquid crystals the director field $\mathbf{n}$ and the velocity field $\mathbf{v}$ are coupled by nonlinear nematohydrodynamic equations – see e.g. [48,91] that in the lowest approximation lead to the same governing equations of viscous fingering as for isotropic fluids with some effective viscosity μ_{eff} and effective surface tension σ_{eff} [92,93].

For a $\mathbf{v} = v_x(z)$ velocity profile one has to distinguish different orientations of the nematic director which give different μ_{eff} as follows:

i. For planar cells and $\mathbf{n} \parallel x$ ($n_x \approx 1, n_y \ll 1, n_z \ll 1$): $\mu_{eff} = (\alpha_3+\alpha_4+\alpha_6)/2 = \eta_1$

ii. For homeotropic cells ($\mathbf{n} \parallel z$, i.e. $n_x \ll 1, n_y \ll 1, n_z \approx 1$): $\mu_{eff} = (\alpha_4 + \alpha_5 - \alpha_2)/2 = \eta_2$

iii. For planar cells and $\mathbf{n} \parallel y$ ($n_x \ll 1, n_y \approx 1, n_z \ll 1$): $\mu_{eff} = \alpha_4/2 = \eta_3$.

Here α_i ($i = 1-6$) are the Leslie viscosity coefficients and η_i ($i = 1, 2, 3$) are the Miezowitz viscosity coefficients (see e.g. [48]) that satisfy relation: $\eta_1 < \eta_3 < \eta_2$.

The orientation of the director with respect to the interface defines the surface tension of the air – nematic interface in the range of: $\sigma_\perp \leq \sigma_{eff} \leq \sigma_\parallel$, where $\sigma_\parallel$ and $\sigma_\perp$ are the surface tension parallel and perpendicular to $\mathbf{n}$, respectively. However, the change in the surface tension from $\sigma_\parallel$ to $\sigma_\perp$ is typically in the range of 20-50% for nematic substances [94], while the anisotropy in the viscosity is much larger: typically $\eta_2 \simeq 5\eta_1$ far below the nematic $\rightarrow$ isotropic phase transition temperature – see e.g. [95,96].

A number of experiments have been done in systems with a nematic liquid crystal as the more viscous fluid, for review see [67,92,93,97]. Four basic types of morphologies have been found in these systems: tip-splitting branches (SB), dense branching (DB), sparse dendritic (SD) and dense dendritic (DD) – see Fig. 6 [90,98]. The tip stabilization and the appearance of the dendritic pattern in a certain pressure – temperature range has been attributed to the anisotropy of μ_{eff} at the tip, induced by flow alignment of the nematic [88,99]. Recently, it has been shown by numerical simulation that non-Newtonian behavior can suppress tip splitting and produce dendritic growth with side-branches [100,101].

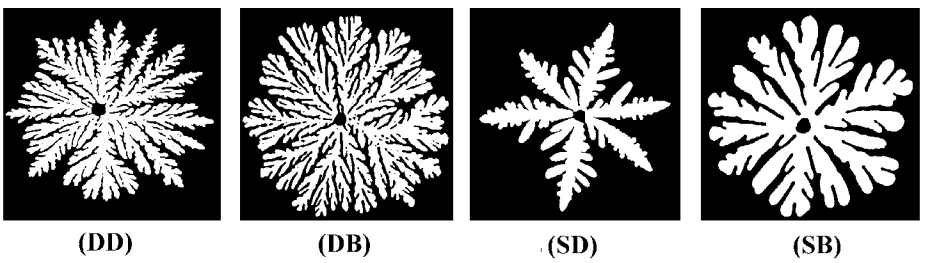

Fig. 6. Four basic morphologies of the air – 8CB liquid crystal interface.

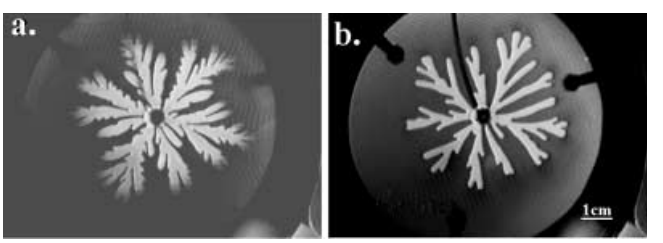

Fig. 7. Morphological transition at the nematic – air interface induced by electric field ($p_e = 123$ mbar); (a) $E = 0$, (b) $E = 0.5$ V/μm.

The effective viscosity μ_{eff} of the nematics and σ_{eff} of the N – air interface in the shear plane can be tuned by an electric field. Namely, using nematic liquid crystals with positive dielectric anisotropy $\varepsilon_a = \varepsilon_{\parallel} - \varepsilon_{\perp} > 0$ and applying an external electric field E perpendicular to the plane of the cell, one has two limiting cases: **(a)** When the shear torque [48] exerted on the director is much larger than the electric torque (high excess pressure p_e and low E, elastic torques are negligible in our case) the effective viscosity is $\mu_{eff} \approx \eta_1$ and the surface tension is $\sigma_{eff} \approx \sigma_{\parallel}$ (assuming 'flow alignment' so that $\mathbf{n}$ is in the plane of the cell and perpendicular to the interface). **(b)** In the opposite case, when the electric torque is much larger than the shear torque, we have $\mu_{eff} \approx \eta_2$ and

$\sigma_{eff} \approx \sigma_{\perp}$ (**n** is perpendicular to the bounding glass plates). Consequently, in the experiments μ_{eff} and σ_{eff} can be tuned depending on E and p_e in the range of: $\eta_1 \leq \mu_{eff} \leq \eta_2$ and $\sigma_{\perp} \leq \sigma_{eff} \leq \sigma_{\parallel}$ that can induce a morphological transition as it is shown in Fig. 7 [102].

Viscous fingering phenomena at the SmA – air interface are less well studied compared to the N – air system. Tip splitting of the fingers growing in the homeotropic SmA phase of 8CB has been observed leading to the dense-branching morphology (no anisotropy in the flow in the shear plane) with a decrease of the characteristic finger width with decreasing temperature [99]. In case of planar SmA phase (where anisotropy in the flow is present in the shear plane) however, a change in the direction of 'easy growth' of the fingers has been detected as a function of the excess pressure p_e [98] – a phenomenon that is still not fully understood.

Acknowledgement

We wish to thank S. Kai, L. Kramer, A. Krekhov, A. Hernández-Machado, J. Casademunt and R. Folch for discussions. Support of the Hungarian Research Grants No. OTKA-F022771, OTKA-T022772, OTKA-T031808 and the EU TMR Research Network 'Patterns, Noise and Chaos' is gratefully acknowledged. A.B. thanks the XVII Sitges Conference organizers for the invitation.

References

1. P. G. de Gennes and J. Prost: *The Physics of Liquid Crystals* (Clarendon Press, Oxford, 1993); S. Chandrasekhar, *Liquid Crystals* (Cambridge University Press, 1992)
2. G. Gray and J. Goodby: *Smectic Liquid Crystals* (Leonard Hill, Glasgow and London, 1984)
3. L. M. Blinov, V. G. Chigrinov: *Electrooptic Effects in Liquid Crystal Materials*, (Springer, Berlin, 1994).
4. Á. Buka, L. Kramer (Eds.): *Pattern Formation in Liquid Crystals*, (Springer, Berlin, 1995).
5. L. Kramer, W. Pesch: 'Electrohydrodynamic Instabilities in Nematic Liquid Crystals'. In: *Pattern Formation in Liquid Crystals*. ed. by Á. Buka, L. Kramer, (Springer, Berlin, 1995) pp. 221–255.
6. R. Williams: J. Chem. Phys. **39**, 384 (1963)
7. E. Bodenschatz, W. Zimmermann, L. Kramer: J. Phys. (France) **49**, 1875 (1988)
8. H. Bohatsch, R. Stannarius: Phys. Rev. E **60**, 5591 (1999)
9. S. Kai, K. Yamaguchi, K. Hirakawa: Jpn. J. Apl. Phys. **14**, 1653 (1975)
10. A. Hertrich, W. Decker, W. Pesch, L. Kramer: J. Phys. (France) II **2**, 1915 (1992); L. Kramer, A. Hertrich, W. Pesch: 'Electrohydrodynamic convection in nematics: the homeotropic case.' In: *Pattern Formation in Complex Dissipative Systems*. ed. by S. Kai. (World Scientific, Singapore, 1992).
11. S. Kai, K. Hayashi, Y. Hidaka: J. Phys. Chem. **100**, 19007 (1996)

12. Y. Hidaka, J.-H. Huh, K. Hayashi, S. Kai, M. I. Tribelsky: Phys. Rev. E **56**, R6256 (1997); Y. Hidaka, J.-H. Huh, K. Hayashi, M. I. Tribelsky, S. Kai: J. Phys. Soc. Jpn. **66**, 3329 (1997)
13. J.-H. Huh, Y. Hidaka, S. Kai: J. Phys. Soc. Jpn. **67**, 1948 (1998)
14. P. Tóth, Á. Buka, J. Peinke, L. Kramer: Phys. Rev. E **58**, 1983 (1998)
15. J.-H. Huh, Y. Hidaka, S. Kai: Mol. Cryst. Liq. Cryst. **328**, 497 (1999); J.-H. Huh, Y. Hidaka, S. Kai: J. Phys. Soc. Jpn. **68**, 1567 (1999)
16. A. G. Rossberg, N. Éber, Á. Buka, L. Kramer: Phys. Rev E **61**, R25 (2000); N. Éber, A. G. Rossberg, Á. Buka, L. Kramer: Mol. Cryst. Liq. Cryst. in press
17. Á. Buka, P. Tóth, N. Éber, L. Kramer: Phys. Reports, in press
18. H. Richter, Á. Buka, I. Rehberg: Mol. Cryst. Liq. Cryst. **251**, 181 (1994); H. Richter, Á. Buka, I. Rehberg: 'The Electrohydrodynamic Instability in a Homeotropically Aligned Nematic: Experimental Results'. In: *Spatio-Temporal Patterns*, ed. by P. E. Cladis, P. Palffy-Muhoray (Addison-Wesley, 1995) pp. 343–352
19. J.-H. Huh, Y. Hidaka, S. Kai: Phys. Rev. E **58**, 7355 (1998)
20. E. Plaut, W. Decker, A. G. Rossberg, L. Kramer, W. Pesch, A. Belaidi, R. Ribotta: Phys. Rev. Lett. **79**, 2367 (1997)
21. S. Rudroff, H. Zhao, L. Kramer, I. Rehberg: Phys. Rev. Lett. **81**, 4144 (1998); S. Rudroff, V. Frette, I. Rehberg: Phys. Rev. E **59**, 1814 (1999)
22. H. Amm, M. Grigutsch, R. Stannarius: Mol. Cryst. Liq. Cryst. **320** 11 (1998)
23. W. Pesch, L. Kramer: 'General Mathematical Description of Pattern-Formming Instabilities'. In: *Pattern Formation in Liquid Crystals*. ed. by Á. Buka, L. Kramer, (Springer, Berlin, 1995) pp. 69–90.
24. A. G. Rossberg, A. Hertich, L. Kramer, W. Pesch: Phys. Rev. Lett. **76**, 4729 (1996)
25. E. Plaut, R. Ribotta: Phys. Rev. E **56**, R2375 (1997)
26. E. Bodenschatz, W. Pesch, L. Kramer: Physica D **32**, 135 (1988); L. Kramer, E. Bodenschatz, W. Pesch: Comment in Phys. Rev. Lett. **64**, 2588 (1990)
27. S. Nasuno, S. Takeuchi, Y. Sawada: Phys. Rev. A **40**, 3457 (1989); S. Rasenat, V. Steinberg, I. Rehberg: Phys. Rev. A **42**, 5998 (1990)
28. S. Kai, N. Chizumi, M. Kohno: Phys. Rev. A **40**, 6554 (1989)
29. P. Tóth et al.: to be published
30. Orsay Liquid Crystal Group: Mol. Cryst. Liq. Cryst. **12**, 918 (1971)
31. K. Hirakawa, S. Kai: Mol. Cryst. Liq. Cryst. **40**, 261 (1977)
32. S. Kai, Y. Adachi, S. Nasuno: 'Stability Diagram, Defect Turbulence, and New Patterns in Electroconvection in Nematics'. In: *Spatio-Temporal Patterns*, ed. by P. E. Cladis, P. Palffy-Muhoray (Addison-Wesley, 1995) pp. 313–329
33. M. Scheuring, L. Kramer, J. Peinke: Phys. Rev. E **58**, 2018 (1998); H. Amm, R. Stannarius, A. G. Rossberg: Physica D **126** 171 (1999)
34. A. G. Rossberg, L. Kramer: Physica D **115**, 19 (1998)
35. J.-H. Huh, Y. Hidaka, A. G. Rossberg, S. Kai: Phys. Rev. E **61**, 2769 (2000)
36. J.-H. Huh: Electrohydrodynamic Instability and Nonlinear Dynamics in Liquid Crystals with Continuous Rotationaal Symmetry. Ph.D. Thesis. Kyushu University, Fukuoka, 2000
37. P. Petrescu, M. Giurgea: Phys. Lett. **59A**, 41 (1976); M. I. Barnik, L. M. Blinov, M. F. Grebenkin, A. N. Trufanov: Mol. Cryst. Liq. Cryst. **37** 47 (1976); L. Nasta, A. Lupu, M. Giurgea: Mol. Cryst. Liq. Cryst. **71**, 65 (1981)
38. A. N. Trufanov, L. M. Blinov, M. I. Barnik: 'A Novel Type of the Electrohydrodynamic Instability in Nematic Liquid Crystals'. In: *Advances in Liquid Crystal*

Research and Applications, ed. by L. Bata (Pergamon Press, Oxford - Akadémiai Kiadó, Budapest, 1980) pp. 549–560
39. J.-H. Huh, Y. Hidaka, S. Kai, Y. Yusril, N. Éber, T. Tóth-Katona, Á. Buka, to be published
40. R. Ribotta, G. Durand: J. Phys. (France) Coll. C3 **40** C3-334 (1979)
41. A. G. Rossberg: private communication
42. N. Felici: Rev. Gen. Electr. **78**, 717 (1969)
43. L. Kramer, H. Zhao, R. Benitez: private communication
44. C. Fradin, P. L. Finn, H. R. Brand, P. E. Cladis: Phys. Rev. Lett. **81**, 2902 (1998)
45. M. Dennin, G. Ahlers, D.S. Canell: Phys. Rev. Lett. **77**, 2475 (1996)
46. H.R. Brand, C. Fradin, P.L. Finn, W. Pesch, P.E. Cladis: Phys. Lett. A **235**, 508 (1997)
47. Ch. Gähwiller: Phys. Rev. Lett. **28**, 1554 (1972); P. Pieranski, E. Guyon: Comm. on Physics **1**, 45 (1976); M. Cohen, P. Pieranski, E. Guyon, C.D. Mitescu: Mol. Cryst. Lyq. Cryst. **38** 97 (1977); I. Zuniga, F.M. Leslie: Liq. Cryst.**5**, 725 (1989); Europhys. Lett. **9**, 689 (1989)
48. E. Dubois-Violette, P. Manneville: 'Flow Instabilities in Nematics'. In: *Pattern Formation in Liquid Crystals*, ed. by A. Buka, L. Kramer (Springer-Verlag, New York, 1996) pp. 91–163
49. P. Pieranski, E. Guyon: Solid State Comm. **13**, 435 (1973); Phys. Rev. A **9**, 404 (1974); P. Manneville, E. Dubois-Violette: J. Phys. **37**, 285 (1976)
50. M.G. Clark, F.C. Saunders, I.A. Shanks, F.M. Leslie: Mol. Cryst. Liq. Cryst. **70**, 195 (1981); W.R. Burghardt: J. Rheol. **35**, 49 (1991)
51. A.P. Krekhov, L. Kramer, Á. Buka, A.N. Chuvyrov: J. Phys. II (France) **3**, 1387 (1993)
52. A.P. Krekhov and L. Kramer: J. Phys. II (France) **4**, 677 (1994); Phys. Rev. E **53**, 4925 (1996)
53. F. Scudieri: J. Appl. Phys.**49**, 1289 (1978); E. Guazzelli: Thèse de 3me cycle. Université Paris-Sud, Orsay, 1981
54. T. Börzsönyi, A.P. Krekhov, Á. Buka, L. Kramer: Phys. Rev. E **58**, 7419 (1998)
55. S.J. Hogan, T. Mullin, P. Woodford: Proc. R. Soc. London, Ser.A **441**, 559 (1993); R. Linsey, S. Hogan: Preprint (1996).
56. E.N. Kozhevnikov: Zh. Eksp. Teor. Fiz. **91**, 1346 (1986); Sov. Phys. JETP **64**, 793 (1986)
57. P. Pieranski, E. Guyon: Phys. Rev. Lett. **39**, 1281 (1977); E. Guazzelli, E. Guyon: J. Phys. (France) **43**, 985 (1982); E. Guazzelli: In: *Nematics, NATO ASI Vol.332*. Kuwer, Dordrecht, 1991; E. Guazzelli, G. Dewel, P. Borckmans, D. Walgraef: Physica D **35**, 220 (1989); E. Dubois-Violette, F. Rothen: J. Phys. (France) **39**, 1039 (1978)
58. J.-M. Dreyfus, P. Pieranski: J. Phys. (France) **42**, 459 (1981).
59. T. Börzsönyi, Á. Buka, A.P. Krekhov, O.A. Scaldin, L. Kramer: Phys. Rev. Lett. **84**, 1934 (2000)
60. A. de Wit, Adv. Chem. Physics, ed. by: I. Prigogine, S.A. Rice (John-Wiley, New York, 1999) Vol.109, p. 435.
61. T. Börzsönyi, Ph.D. dissertation, Budapest, 1998 (in English).
62. A. N. Chuvyrov: Zh. Eksp. Teor. Fiz. **82**, 761 (1982) [Sov. Phys. JETP **55**, 451 (1982); A.N. Chuvyrov, O.A. Scaldin, V.A. Delev: Mol. Cryst. Liq. Cryst.**215**, 187 (1992).
63. F. Scudieri: Ann. Phys.**3**, 311 (1978).
64. *Solids far from equilibrium*. Ed. by C. Godrèche (Cambridge University Press, Cambridge, 1992)

65. J. Langer: Rev. Mode. Phys. **52**, 1 (1980)
66. E. Rubinstein, M. Glicksman: J. of Crystal Growth **112**, 97 (1991); K. Koo, R. Ananth, W. Gill: Phys. Rev. A **44**, 3782 (1991); M. Muschol, D. Liu, H. Cummins: Phys.Rev. A **46**, 1038 (1992); M. Glicksmann, N. Singh: J. Cryst. Growth **98**, 277 (1989); A. Dougherty, J. Gollub: Phys. Rev. A **38**, 3043 (1988)
67. K. McCloud, J. Maher: Phys. Rep. **260**, 139 (1995)
68. D. Grier, E. Ben-Jacob, R. Clarke, L. Sander: Phys. Rev. Lett. **56**, 1264 (1986); D. Barkey, F. Oberholtzer, Q. Wu: Phys. Rev. Lett. **75**, 2980 (1995)
69. E. Ben-Jacob, O. Shochet, A. Tenenbaum, I. Cohen, A. Czirok, T. Vicsek: Fractals **2**, 15 (1994)
70. O. Shochet, K. Kassner, E. Ben-Jacob, S. Lipson, H. Müller-Krumbhaar: Physica A **181**, 136 (1992)
71. H. Müller-Krumbhaar: 'Theory of dendritic crystal growth'. In: *Morphology of crystals* ed. by: I. Sunagawa (Terra Scientific Publishing Company, Tokyo, 1987)
72. O. Penrose, P. C. Fife: Physica D **43**, 44 (1990); S.-L. Wang, R. F. Sekerka, A. A. Wheeler, B. T. Murray, S. R. Coriell, R. J. Braun, G. B. McFadden: Physica D **69**, 189 (1993); A. Karma, W. J. Rappel: Phys. Rev. E **57**, 4323 (1998); R. Kobayashi: Physica D **63**, 410 (1993)
73. R. Pratibha, N. Madhusudana: J. Phys. II. (France) **2**, 383 (1992); H. Naito, M. Okuda, O.-Y. Zhong-can: Phys. Rev. Lett. **70**, 2912 (1993)
74. P. Oswald, J. Malthéte, P. Pelcé: J. Phys. **50**, 2121 (1989); J. Géminard, P. Oswald, D. Temkin, J. Malthête: Europhys.Lett. **22**, 69 (1993)
75. P. Oswald: J. Phys. (France) **49**, 2119 (1988); P. Oswald: J. Phys. (France) **49**, 1083 (1988); J. Géminard, P. Oswald: J. Phys. II (France) **4**, 959 (1994)
76. P. Oswald, J. Bechhoefer, A. Libchaber: Phys. Rev. Lett. **58**, 2318 (1987)
77. J. Bechhoefer, P. Oswald, A. Libhaber, C. Germain: Phys.Rev. A **37**, 1691 (1988)
78. J. Hutter, J. Bechhoefer: Physica A **239**, 103 (1997)
79. T. Tóth Katona, T. Börzsönyi, Z. Váradi, J. Szabon, Á. Buka, R. González-Cinca, L. Ramírez-Piscina, J. Casademunt, A. Hernández-Machado: Phys. Rev. E. **54**, 1574-1583 (1996)
80. R. González-Cinca, L. Ramírez-Piscina, J. Casademunt, A. Hernández-Machado, L. Kramer, T. Tóth Katona, T. Börzsönyi, Á. Buka: Physica D **99**, 359 (1996)
81. T. Börzsönyi, Á. Buka, L. Kramer: Phys. Rev. E **58**, 6236 (1998)
82. R. González-Cinca, L. Ramírez-Piscina, J. Casademunt, A.Hernández-Machado, T. Tóth Katona, T. Börzsönyi, Á. Buka: J. Cryst. Growth **193**, 712 (1998)
83. T. Börzsönyi, T. Tóth-Katona , Á. Buka, L. Gránásy: Phys. Rev. Lett. **83**, 2853 (1999); T. Börzsönyi, T. Tóth-Katona , Á. Buka, L. Gránásy: Phys. Rev. E **62**(September 2000)
84. A.A. Chernov: Contemporary Physics **30**, 251 (1989); W.K. Burton, N. Cabrera, F.C. Franck: Phil. Trans. Roy. Soc. London **2431**, 40 (1951)
85. T. Börzsönyi, S. Akamatsu G. Faivre: To be published.
86. D. Bensimon, L. Kadanoff, S. Liang, B. Shraiman, C. Tang: Rev. Mod. Phys. **58**, 977 (1986)
87. L. Paterson: J. Fluid Mech. **113**, 513 (1981)
88. Á. Buka, P. Palffy-Muhoray: Phys. Rev. A **36**, 1527 (1987)
89. K. Makino, M. Kawaguchi, K. Aoyama, T. Kato: Phys. Fluids **7**, 455 (1995); M. Kawaguchi, K. Makino, T. Kato: Physica A **246**, 385 (1997)
90. Á. Buka, J. Kertész, T. Vicsek: Nature **323**, 424 (1986)
91. H. Pleiner, H. Brand: 'Hydrodynamics and Electrohydrodynamics of Liquid Crystals'. In: *Pattern Formation in Liquid Crystals*, ed. by Á. Buka, L. Kramer (Springer-Verlag, New York, 1996) pp. 15–67

92. L. Lam, H. Morris, R. Shao, S. Yang, Z. Liang, S. Zheng, H. Liu: Liq. Cryst. **5**, 1813 (1989)
93. Á. Buka: 'Viscous Fingering in Liquid Crystals'. In: *Pattern Formation in Liquid Crystals*, ed. by Á. Buka, L. Kramer (Springer-Verlag, New York, 1996) pp.291–305
94. V. Tsvetkov, O. Tsvetkov, V. Balandin: Mol. Cryst. Liq. Cryst. **329**, 305 (1999)
95. H. Kelker, R. Hatz: *Handbook of Liquid Crystals*. (Verlag Chemie, Weinheim 1980)
96. C. Gähwiller: Mol. Cryst. Liq. Cryst. **20**, 301 (1973)
97. A. Sonin: Rivista del Nuovo Cimento **14**, 1 (1991)
98. Á. Buka, P. Palffy-Muhoray: J. Phys. (France) **49**, 1319 (1988)
99. Á. Buka, P. Palffy-Muhoray, Z. Rácz: Phys. Rev. A **36**, 3984 (1987)
100. K. Kondic, M. Shelley, P. Palffy-Muhoray: Phys. Rev. Lett. **80**, 1433 (1998)
101. R. Folch, J. Casademunt, A. Hernández-Machado: Phys. Rev. E **61**, 6632 (2000)
102. R. Folch, T. Tóth-Katona, Á. Buka, J. Casademunt, A. Hernández-Machado: To be published

Inviscid Two-Dimensional Fluid Dynamics Experiments with Magnetized Electron Columns

Joel Fajans and Daniel Durkin

Physics Dept, U.C. Berkeley, Berkeley CA 94720-7300, USA

1 Introduction

Inviscid two-dimensional (2D) fluid phenomena like hurricanes, jet streams, Jupiter's Red Spot, and protoplanetary disks are quite common in Nature. Unfortunately, 2D fluid phenomena are difficult to study in the laboratory because the finite size of laboratory apparati commonly introduces unwanted viscous and friction effects. Magnetized electron columns, however, behave like two-dimensional fluids and are not subject to viscous effects [1]. The columns behave like fluids because the equations which govern their behavior, the Drift-Poisson Equations, are identical to the 2D inviscid Euler equations which govern 2D fluids. More precisely, the columns follow $\boldsymbol{E} \times \boldsymbol{B}$ dynamics, i.e their velocity $\boldsymbol{v}$ is

$$\boldsymbol{v} = \boldsymbol{E} \times \boldsymbol{B}/B^2, \tag{1}$$

where $\boldsymbol{E}$ is the self-consistent electric field from the columns and any external potentials, and $\boldsymbol{B}$ is a strong external magnetic field. (The columns are trapped in a Malmberg-Penning trap; see Fig. 1.) The electric field can be derived from a potential, namely $\boldsymbol{E} = -\nabla\phi$. The vorticity is defined just as it would be for a fluid: $\Omega = \nabla \times \boldsymbol{v}$. Expressing the velocity in terms of the potential and using Poisson's equation yields

$$\Omega = \frac{e}{\epsilon_0 B} n, \tag{2}$$

where n is the electron density; thus vorticity is proportional to density. By creating electron columns with the appropriate density, we can model 2D fluid flows. To aid in the creation of arbitrary initial conditions, we have developed a photocathode source [2] that allows us to create any desired initial condition simply by projecting an appropriate light image onto the photocathode.

2 Decay of Turbulence

Magnetized plasma columns can be used to study the decay of 2D turbulence. For example, Fig. 2 shows the decay of a complicated initial vorticity distribution to a simple final state. The relaxation involves subprocesses like advection, instabilities, filamentation, merger, vortex in vortex interactions, and sometimes cooling into vortex crystals. Examples of each of these process will be given in this paper and more information can be found in the cited literature.

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 319–327, 2001.
© Springer-Verlag Berlin Heidelberg 2001

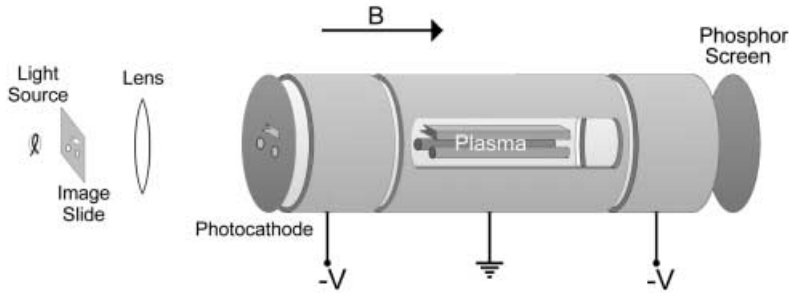

Fig. 1. Malmberg-Penning trap geometry. Negatively biasing the end cylinders provides axial confinement, and the axial magnetic field provides radial confinement. The electron columns (plasma) are emitted from the photocathode on the left, and loaded into the trap by temporarily grounding the left cylinder. The columns are diagnosed by grounding the right cylinder, thereby allowing the columns to stream out onto the phosphor screen. The images thus created are detected by a CCD camera. Movies of the evolution of the columns are obtained by successively regenerating the initial configuration and grounding the right cylinder at successively later times. Details of the trap operation can be found in Refs. [2,3].

2.1 Advection

Two vortices will interact with each other due to their mutual electric fields/stream functions. If the two vortices have the same sign, as in Fig. 3, the $\boldsymbol{E}$ fields from the two vortices will be radial but point in opposite directions. The resulting $\boldsymbol{E} \times \boldsymbol{B}$ velocities will be perpendicular to the $\boldsymbol{E}$ fields and point oppositely. As the vortices move, the $\boldsymbol{E}$ field directions will adjust; the end result is that the vortices will orbit around each other in a circle. If the two vortices have the opposite sign, the $\boldsymbol{E}$ fields will point in the same direction, and the vortices will drift together in the direction perpendicular to the line separating the vortices. For example, a vortex near a wall will move under the influence of its oppositely-signed image vortex; since the line connecting the original vortex and its image is radially directed, the drift direction will be azimuthal. Thus the vortex and its image will drift azimuthally, revolving around the trap wall [4]. Obviously electrons all have the same sign, so one might think that, aside from image vortices, magnetized electron columns would all have the same sign of vorticity, and it would be impossible to generate opposite sign vortices – anti-vortices. However, superimposing a uniform background density on a vorticity distribution is equivalent to transforming to a rotating frame, and such a transformation does not change the dynamics. In this rotating frame, the holes, i.e. the regions where the density is below the density of the uniform background, act exactly like anti-vortices. This point is illustrated in Fig. 4 which shows two holes revolving around each other until they merge. (Merger is discussed in Sec. 2.3.)

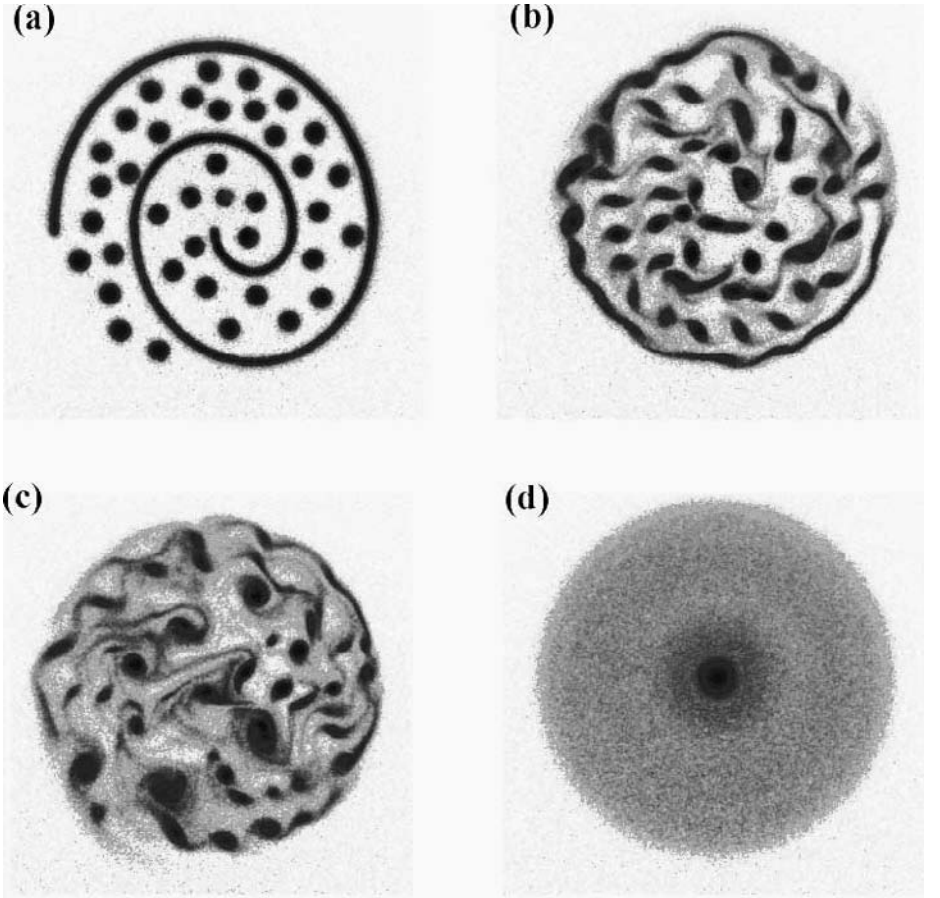

Fig. 2. An initially complicated state (a), evolving through successively simpler states (b-c) to a final relaxed state (d).

2.2 Filamentation

Several processes will induce filamentation. For example, a sufficiently elliptical vortex will spawn filaments from its tips (Fig. 5) [5]. More importantly, a strong vortex can rip apart a weak vortex (Fig. 6) [6]. Filamentation is important because it introduces fine scale structure that coarse graining, or the small, but unavoidable viscosity in the system will eventually smooth out. Typically this creates a tenuous background density/vorticity.

2.3 Merger

When two vortices of the same sign are sufficiently close to each other, they will merge into one larger vortex and several filaments. (See Fig. 4.) For identical

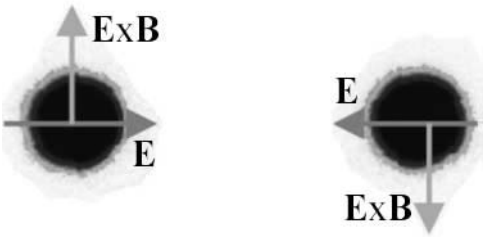

Fig. 3. Two vortices of the same sign. The left electric field arrow is the field from the right vortex, and vice versa. As the two electric fields are equal and opposite, the $\boldsymbol{E} \times \boldsymbol{B}$ velocities point azimuthally, and the two vortices orbit each other.

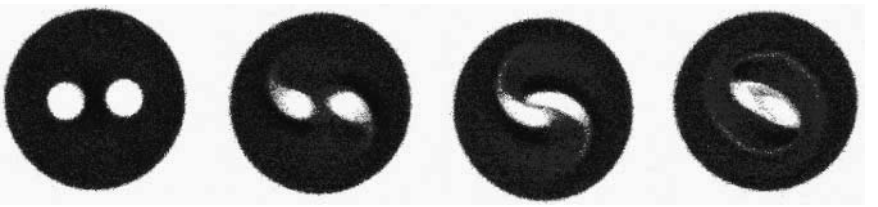

Fig. 4. The merger of two holes/anti-vortices.

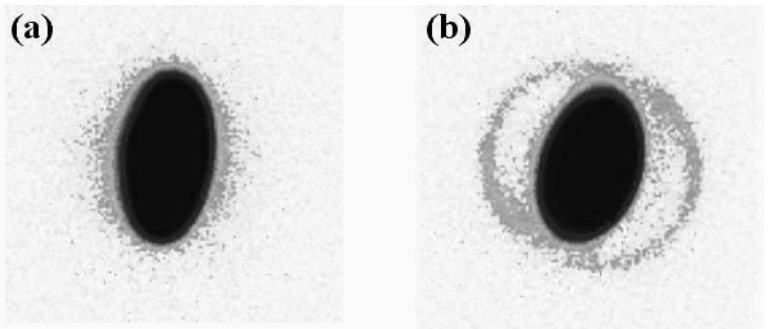

Fig. 5. An ellipse will spawn spiral filaments from its tips.

vortices, the precise merger criterion is that the centers of the two vortices must not be separated by more than 3.4 times the radius of the vortices. This process is very important to the decay of turbulence as it is the major mechanism by which discrete vortices aggregate. Merger was studied in detail by Fine, et. al [7].

2.4 Kelvin-Helmholtz Instabilities

A hollow ring of density/vorticity will be unstable to the Kelvin-Helmholtz instability (Fig. 7) [1]. The Kelvin-Helmholtz instability is due to shear in the

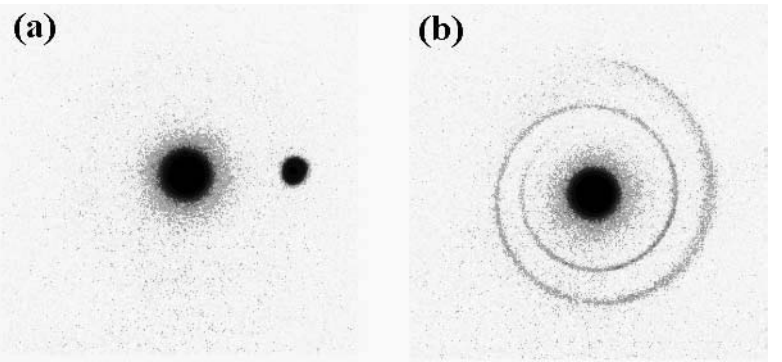

Fig. 6. A strong vortex will tear a weak vortex apart.

velocity; the inner edge of the ring moves slower than the outer edge. The number of lobes that the ring breaks up into is inversely proportional to the width of the ring [8,9]. More generally, any vorticity filament will be subject to this instability if the filament has sufficient shear.

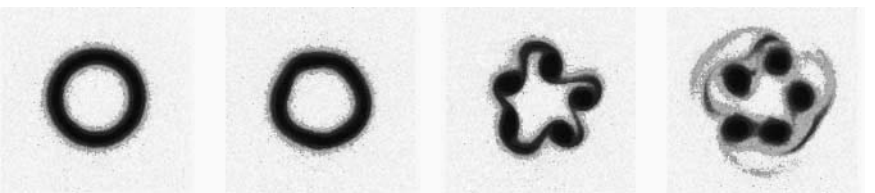

Fig. 7. The breakup of a hollow ring of vorticity/density due to the Kelvin-Helmholtz instability.

2.5 Vortex in Vortex Interactions

When a intense, point-like vortex is placed inside an diffuse, extended vortex, the point-like vortex will often induce a wave on the diffuse vortice's surface (Fig 8). The will grow nonlinearly and eventually break, entraining an anti-vortex. Depending on the parameters, the point vortex may be able to drag the anti-vortex into the extended vortice's interior. This interaction was only discovered very recently, and experimental measurements [10] are in good agreement with theoretical expectations [11]. The interaction's importance stems from the turbulence that it can engender in the initially relaxed extended vortex.

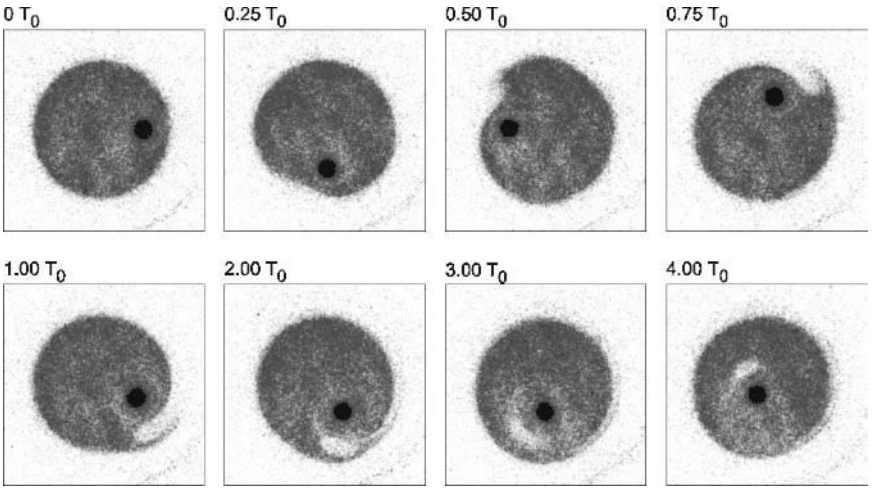

Fig. 8. Vortex in vortex interaction. The time T_0 is the time for the extended vortex to rotate once. The surface wave breaks at approximately $0.5T_0$, entrains an anti-vortex by $2T_0$, and the point vortex pulls the anti-vortex into the extended vortex interior by $3T_0$

2.6 Vortex Cooling and Crystallization

Interactions between initially randomly arranged vortices and an extended, diffuse background vortex will sometimes lead to the vortices cooling into a crystal [12,13,14,15]. A typical sequence is illustrated in Fig. 9, and a complex crystal is shown in Fig. 10 [16]. Figure 11 shows the measured "jitter" as a function of

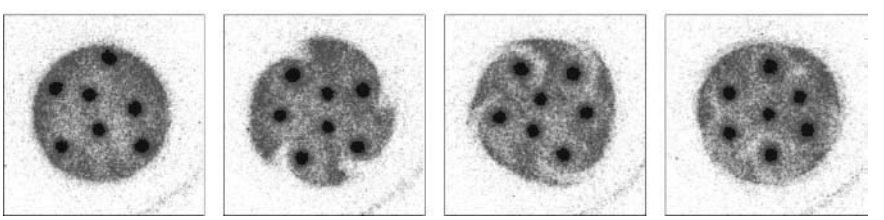

Fig. 9. An initially random configuration of seven vortices in a background cooling into a crystal.

time for a configuration of four randomly placed vortices in a background. The cooling results from the stirring of the background by the vortices. This maximizes the entropy of the background [19], producing a final state that favors crystallization. Computer simulations indicate that the cooling rate depends on the background density, with the fastest rates occuring for densities which are neither too high or too low [15]. However, typical computer simulations take

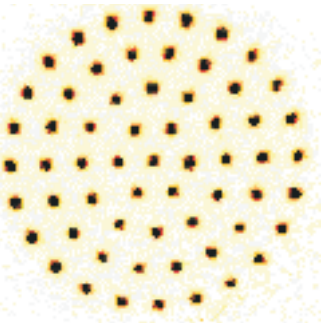

Fig. 10. A crystal with 61 vortices. This triangular lattice crystal is the fourth in a series of "magic" numbers, the first crystals having 7, 19 and 37 vortices. In this case, the crystal was created directly by the photocathode rather than by cooling from an initially random state. Vortex crystals have a long history; Lord Kelvin and J.J Thomson postulated that molecules might be stable configurations of vortices [17,18].

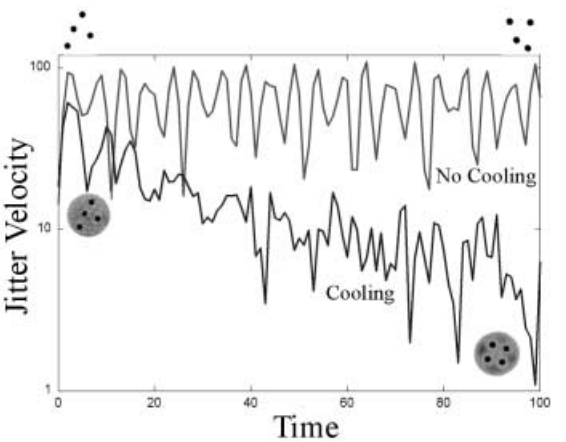

Fig. 11. Time vs. "Jitter Velocity" for two configuration: one with no background which does not cool, and one that is identical except for the addition of a background, which does cool. The Jitter Velocity is the velocity by which a configuration deviates from a uniform rotation.

many days, and only a limited number of initial conditions were explored. Our experimental measurements show that the cooling rate is at least as dependant on the closeness of the vortices to the background edge as it is on the background density.

2.7 Vortex Cooling and Unstable Equilibria

Sometimes the cooling process will take the system near an unstable equilibria; i.e. a crystal configuration that is unstable (Fig. 12). These unstable equilibria

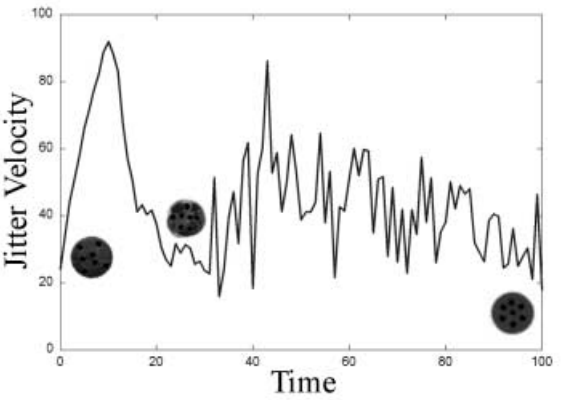

Fig. 12. Time vs. Jitter Velocity for the sequence of Fig. 9. Note how the system temporarily settles into what turns out to be an unstable equilibrium near a time of 25. The unstable equilibrium configuration is illustrated by the middle small inset figure; the final stable configuration is illustrated by the inset figure on the right.

were recently predicted by Aref and Vainchtein [20] who correctly assert that the system will stagnate in the vicinity of an unstable equilibrium.

3 Conclusions

Magnetized plasma columns are a powerful tool to study 2D fluid dynamics. In conjunction with analytic work and computer simulations, magnetized electron columns can be used to study many theoretical issues as well as problems in Nature. Our recently developed photocathode source, which allows us to create arbitrary initial conditions, greatly extends the range of experiments that can be undertaken.

Acknowledgments

Many people contributed to this work at Berkeley, including E.Yu. Backhaus, R. Chu, E. Gilson, J. Notte, A.J. Peurrung, J. Siegel, J.S. Wurtele, and L. Zimmerman. Experimentalists cited at other institutions include C.F. Driscoll and K. Fine at UCSD, and D.L. Eggleston at Occidental. This work was supported by the Office of Navel Research.

References

1. R. H. Levy, Phys. Fluids **8**, 1288 (1965).
2. D. Durkin and J. Fajans, Rev. Sci. Instrum. **70**, 4539 (1999).

3. J. H. Malmberg, C. F. Driscoll, B. Beck, D. L. Eggleston, J. Fajans, K. Fine, X. P. Huang, and A. W. Hyatt, in *Nonneutral Plasma Physics*, edited by C. Roberson and C. Driscoll (American Institute of Physics, New York, 1988), Vol. AIP 175, p. 28.
4. W. D. White, J. H. Malmberg, and C. F. Driscoll, Phys. Rev. Lett. **49**, 1822 (1982).
5. M. V. Melander, N. J. Zabusky, and J. C. McWilliams, J. Fluid Mech. **195**, 303 (1988).
6. D. L. Eggleston, Phys. Plasmas **1**, 3850 (1995).
7. K. S. Fine, C. F. Driscoll, J. H. Malmberg, and T. B. Mitchell, Phys. Rev. Lett. **67**, 588 (1991).
8. A. J. Peurrung and J. Fajans, Phys. Fluids A **5**, 493 (1993).
9. A. J. Peurrung, J. Notte, and J. Fajans, J. Fluid Mech. **252**, 713 (1993).
10. D. Durkin and J. Fajans, submitted to Phys. Rev. Lett.
11. D.Z. Jin and D.H.E. Dubin, submitted to Phys. Plasmas.
12. T. Havelock, Philos. Mag. **11**, 617 (1931).
13. L. Campbell and R. Ziff, A Catalog of Two-Dimensional Vortex Patterns, Los Alamos Scientific Laboratory Report No. LA-7384-MS, 1978.
14. K. Fine, A. Cass, W. Flynn, and C. Driscoll, Phys. Rev. Lett. **75**, 3277 (1995).
15. D. A. Schecter, D. H. E. Dubin, K. S. Fine, and C. F. Driscoll, Phys. Fluids **11**, 905 (1999).
16. D. Durkin and J. Fajans, Phys. Fluids **12**, 298 (2000).
17. W. T. Kelvin, *Mathematical and Physical Papers, iv.* (Cambridge University Press, Cambridge, 1878), p. 135.
18. J. Thomson, *Treatise on Vortex Rings* (Macmillan, London, 1883), p. 94.
19. D. Z. Jin and D. H. E. Dubin, Phys. Rev. Lett. **80**, 4434 (1998).
20. H. Aref and D. Vainchtein, Nature **392**, 769 (1998).

Interaction of Nearly-Inviscid, Multi-mode Faraday Waves and Mean Flows

María J. Higuera[1], José M. Vega[1], and Edgar Knobloch[2]

[1] E.T.S. Ingenieros Aeronáuticos,
Pza. Cardenal Cisneros 3, 28040 Madrid, Spain
[2] Department of Physics, University of California, Berkeley,
California 94720

1 Introduction

Faraday waves [1] are gravity-capillary waves that are excited on the surface of a fluid when its container is vibrated vertically and the vertical acceleration exceeds a threshold value. These waves have received much attention in the literature both as a basic fluid dynamical problem and as a paradigm of a pattern-forming system [2,3,4]. Unfortunately, in the low viscosity limit, there are several basic issues that remain unresolved, particularly in connection with the generation of mean flows in the bulk. The viscous part of these flows (also called streaming flow or acoustic streaming) is driven by the oscillatory boundary layers attached to the solid walls and the free surface by well-known mechanisms first uncovered by Schlichting [5] and Longuet-Higgins [6]. This mean flow has been shown recently to affect the dynamics of the primary waves at leading order in a related, laterally vibrated system [7]. This is somewhat similar to the effect of an internal circulation on surface wave dynamics in drops [8].

The main object of this note is to present and discuss a set of asymptotically correct (as viscosity and wave steepness go to zero) coupled amplitude-mean flow (CASF) equations for multi-mode Faraday waves in containers of general cross-section. It will be seen that it is asymptotically inconsistent to ignore the streaming flow while retaining the usual cubic nonlinear terms whenever more than one surface mode is present. Symmetries of the problem, if present, limit the scope of the interaction with the mean flow; however, as the number of excited surface modes increases, the interaction becomes stronger and stronger. Despite the fact that the CASF equations include 3-D continuity and Navier-Stokes-like equations for the mean flow some qualitative properties of the solutions can be established. As an example, we consider in §4 two-mode Faraday waves in almost-square, rectangular containers, as in the well-known experiments by Simonelli and Gollub [9] and Feng and Sethna [10].

2 Formulation

We consider a vertically vibrated cylindrical container of general cross-section Σ. In order to avoid the uncertainties associated with the modeling of contact-line dynamics [2] and additional difficulties due to a strong singularity in the

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 328–337, 2001.
© Springer-Verlag Berlin Heidelberg 2001

velocity at a moving contact line when the contact angle differs from 0 or π [11], the contact line is assumed to be pinned to the upper edge of the vertical wall of the container, as done in some Faraday experiments [12] in order to eliminate the lateral meniscus and the associated meniscus waves. For nondimensionalization, we use the unperturbed depth, ℓ, as characteristic length and, in order to take into account simultaneously both gravity and microgravity conditions, we use the *gravity-capillary time* $[g/\ell + \sigma/(\rho\ell^3)]^{-1/2}$ as the characteristic time. Here g is the gravitational acceleration, σ is the surface tension coefficient and ρ is the density, all assumed to be constant. We use a Cartesian coordinate system attached to the vibrating container, with the $z = 0$ plane at the unperturbed free surface, assumed to be horizontal. The governing continuity and momentum conservation equations are

$$\boldsymbol{\nabla}\cdot\boldsymbol{v} = 0, \quad \partial\boldsymbol{v}/\partial t - \boldsymbol{v}\times\boldsymbol{\nabla}\times\boldsymbol{v} = -\boldsymbol{\nabla}p + C_g\Delta\boldsymbol{v} \tag{1}$$

if $(x, y) \in \Sigma$ and $-1 < z < f$, with the boundary conditions

$$\boldsymbol{v} = \boldsymbol{0} \quad \text{if } z = -1 \text{ or if } (x, y) \in \partial\Sigma, \quad f = 0 \quad \text{if } (x, y) \in \partial\Sigma, \tag{2}$$

$$\boldsymbol{v}\cdot\boldsymbol{n} = (\partial f/\partial t)(\boldsymbol{e}_z\cdot\boldsymbol{n}),\ [(\boldsymbol{\nabla}\boldsymbol{v} + \boldsymbol{\nabla}\boldsymbol{v}^\top)\cdot\boldsymbol{n}]\times\boldsymbol{n} = \boldsymbol{0}, \quad \text{if } z = f, \tag{3}$$

$$\begin{aligned} p - |\boldsymbol{v}|^2/2 - (1-S)f + S\boldsymbol{\nabla}\cdot[\boldsymbol{\nabla}f/(1+|\boldsymbol{\nabla}f|^2)^{1/2}] \\ = C_g[(\boldsymbol{\nabla}\boldsymbol{v} + \boldsymbol{\nabla}\boldsymbol{v}^\top)\cdot\boldsymbol{n}]\cdot\boldsymbol{n} + 4\varepsilon\omega^2 f\cos 2\omega t \quad \text{if } z = f, \end{aligned} \tag{4}$$

and appropriate initial conditions. Here p (= pressure + $|\boldsymbol{v}|^2/2 + (1-S)z + 4\varepsilon\omega^2 z\cos 2\omega t$) is a *modified pressure*, $\boldsymbol{v}$ is the *velocity*, f is the *vertical deflection* of the free surface, $\boldsymbol{n}$ is the outward unit normal to the free surface, $\partial\Sigma$ stands for the boundary of the container cross-section Σ, and $\boldsymbol{e}_z$ is the upward unit vector. The real parameters $\varepsilon > 0$ and 2ω are the forcing amplitude and frequency, $C_g = \nu/(g\ell^3 + \sigma\ell/\rho)^{1/2}$ (with ν = kinematic viscosity) is a *capillary-gravity number* and $S = \sigma/(\sigma + \rho g\ell^2)$ is a *gravity-capillary balance parameter*. These are related to the usual *capillary number* $C = \nu\sqrt{\rho/(\sigma\ell)}$ and *Bond number* $B = \rho g\ell^2/\sigma$ as $C_g = C/(1+B)^{1/2}$ and $S = 1/(1+B)$.

Although the CASF equations below apply in a wider parameter range, for simplicity we consider the (nearly-inviscid, nearly-resonant, weakly-nonlinear) distinguished limit

$$|\omega - \Omega| \sim C_g^{1/2} \sim \varepsilon \ll 1, \tag{5}$$

where Ω is an inviscid eigenfrequency of the linearized problem around the quiescent state. We assume that Ω has (algebraic and geometric) multiplicity $N \geq 1$. The situation $N > 1$ is nongeneric in the absence of symmetry in the sense that it is of higher codimension. However, in the presence of symmetry multiple symmetry-related modes can be excited simultaneously. Thus we consider small perturbations that split the eigenfrequency Ω into N independent ones, $\Omega_1,\ldots,\Omega_N$, which will be assumed to be such that

$$|\Omega_k - \Omega| \sim \varepsilon \ll 1 \quad \text{for } k = 1,\ldots,N. \tag{6}$$

The role of the streaming flow can be anticipated by considering the linearized problem around the quiescent state, whose solutions are of the form $(\boldsymbol{u}, p, f) =$

$e^{\lambda t}(\boldsymbol{V}, P, F) + \text{c.c.}$, where $(\boldsymbol{V}, P, F)$ is a nontrivial solution of

$$\boldsymbol{\nabla}\cdot\boldsymbol{V} = 0, \quad \lambda\boldsymbol{V} = -\boldsymbol{\nabla}P + C_g\Delta\boldsymbol{V}, \quad \text{if } (x,y) \in \Sigma, -1 < z < 0, \tag{7}$$

$$\boldsymbol{V} = \boldsymbol{0} \text{ if } z = -1 \text{ or if } (x,y) \in \partial\Sigma, \quad F = 0 \quad \text{if } (x,y) \in \partial\Sigma, \tag{8}$$

$$\boldsymbol{V}\cdot\boldsymbol{e}_z = \lambda F, \quad [(\boldsymbol{\nabla}\boldsymbol{U} + \boldsymbol{\nabla}\boldsymbol{U}^\top)\cdot\boldsymbol{e}_z] \times \boldsymbol{e}_z = \boldsymbol{0},$$
$$P - (1-S)F + S\Delta F = C_g[(\boldsymbol{\nabla}\boldsymbol{V} + \boldsymbol{\nabla}\boldsymbol{V}^\top)\cdot\boldsymbol{e}_z]\cdot\boldsymbol{e}_z \text{ if } z = 0. \tag{9}$$

In the nearly-inviscid limit, $C_g \to 0$, this problem exhibits two kinds of *nearly-marginal* (with a small growth rate) *modes*. Either (i) $\lambda = \pm \mathrm{i}\Omega + O(\sqrt{C_g})$, with $|\Omega| \sim |\boldsymbol{V}| \sim |P| \sim |F| \sim 1$ outside some *viscous boundary layers* of thickness $O(\sqrt{C_g})$ attached to the (solid and free) boundary; these are the well-known *surface* (or *nearly-inviscid*) *modes* associated with the surface waves. Or (ii) λ is real, negative and such that $|\lambda| \sim |P| \sim |F| \sim C_g \ll 1$, $|\boldsymbol{V}| \sim 1$; these are the so-called *hydrodynamic* (or *viscous*) modes, and are the ones associated with the streaming flow. Note that the latter modes are non-oscillatory, exhibit small free-surface deflection (both in contrast to the surface modes) and, like the surface modes, are near-marginal (and thus easily excited). Consequently neither mode type can be ignored a priori in a weakly nonlinear theory; the viscous modes may be ignored only in a strictly linear theory of the free surface motion. However, these modes, whose presence was recognized already by Lamb [13], are usually also ignored in the weakly nonlinear theory. We demonstrate below that this is in general inconsistent.

3 Coupled Amplitude-Mean Flow Equations

In addition to the fast time variable $t \sim 1$, we consider the slow variable

$$T = \varepsilon t. \tag{10}$$

Since $C_g \ll 1$, the surface-waves are nearly-inviscid in the bulk (outside the above-mentioned viscous boundary layers), where we may write

$$\begin{aligned}(\boldsymbol{v}, p, f) \quad &= \varepsilon^{1/2} \mathrm{e}^{\mathrm{i}\omega t} \textstyle\sum_{k=1}^{N} A_k(\boldsymbol{V}_k, P_k, F_k) + \text{c.c.} \\ &+ \varepsilon[(\boldsymbol{v}^s, 0, 0) + \textstyle\sum_{k,l=1}^{N} \bar{A}_k A_l(\boldsymbol{\nabla}H_{kl}, P_{kl}, F_{kl}) + OT] + \ldots .\end{aligned} \tag{11}$$

Here $(\boldsymbol{V}_k, P_k, F_k)$ is the limit as $\Omega_k \to \Omega$ of the inviscid eigenfunction associated with the eigenfrequency Ω_k, A_k is the associated complex amplitude that, together with the streaming flow velocity $\boldsymbol{v}^s$, is independent of the fast time t, overbars and c.c. stand for the complex conjugate, and OT denotes *oscillatory terms*. In addition, we assume that the eigenmodes can be selected such that

$$\int_{-1}^{0}\int_{\Sigma} \boldsymbol{V}_k\cdot\boldsymbol{V}_l d\boldsymbol{x} + \int_{\Sigma}[(1-S)\bar{F}_k F_l + S\boldsymbol{\nabla}F_k\cdot\boldsymbol{\nabla}F_l]dxdy = \delta_{kl}, \tag{12}$$

where δ_{kl} is the Kronecker delta. This assumption is made only to simplify the derivation of the CASF equations and is frequently satisfied in practice. It allows

us to obtain the leading order approximation to the energy of the system as

$$\mathcal{E} = \sum_{k=1}^{N} |A_k|^2. \tag{13}$$

Note that the mean flow velocity has been decomposed into two parts, namely (i) the *inviscid mean flow velocity*, $\sum \bar{A}_k A_l \boldsymbol{\nabla} H_{kl}$, with the potential H_{kl} given by

$$\Delta H_{kl} = 0 \text{ if } (x,y) \in \Sigma \text{ and } -1 < z < 0, \tag{14}$$

$$H_{kl} = 0 \text{ if either } z = -1 \text{ or } (x,y) \in \partial\Sigma, \tag{15}$$

$$\partial H_{kl}/\partial z = -\mathrm{i}\tilde{\boldsymbol{\nabla}}\cdot(\bar{F}_k \tilde{\boldsymbol{V}}_l + F_l \bar{\tilde{\boldsymbol{V}}}_k) \text{ if } z = 0, \tag{16}$$

accounting for the normal component of the mean flow velocity at the unperturbed boundary, and (ii) the *viscous mean flow velocity*, or *streaming flow velocity*, required to be tangent to the unperturbed boundary ($\boldsymbol{v}^s\cdot\boldsymbol{e}_z = 0$ at $z = 0$).

The evolution equations for the complex amplitudes and the streaming flow are obtained from the *solvability conditions* for the equations that result when the expansion (11) is substituted into (1) and into the boundary conditions (for the solution in the bulk) that result from *matching conditions* with the solution in the viscous boundary layers. This process is a natural extension to the multimode problem of the asymptotic derivation in [14]. The resonant oscillatory terms at $O(\varepsilon^{3/2})$ yield

$$\begin{aligned} A_k'(T) = -[\delta_k(1+\mathrm{i}) + \mathrm{i}d_k]A_k + \mathrm{i}\textstyle\sum_{l,m,n=1}^{N} \alpha_{klmn}\bar{A}_l A_m A_n \\ +\mathrm{i}\textstyle\sum_{l=1}^{N} \beta_{kl}\bar{A}_l - \mathrm{i}\Omega \sum_{l=1}^{N} \int_{-1}^{0}\int_{\Sigma} \boldsymbol{v}^s\cdot\boldsymbol{g}_{kl}\, d\boldsymbol{x} A_l \end{aligned} \tag{17}$$

for $k = 1, \ldots, N$, and from the non-oscillatory terms, at orders ε and ε^2, one obtains

$$\boldsymbol{\nabla}\cdot\boldsymbol{v}^s = 0, \tag{18}$$

$$\partial\boldsymbol{v}^s/\partial T - [\boldsymbol{v}^s + \textstyle\sum_{k,l=1}^{N} \bar{A}_k A_l(\boldsymbol{\nabla} H_{kl} - \boldsymbol{g}_{kl})] \times (\boldsymbol{\nabla}\times\boldsymbol{v}^s) = -\boldsymbol{\nabla} q^m + \varepsilon D\Delta\boldsymbol{v}^s \tag{19}$$

for $(x,y) \in \Sigma, -1 < z < 0$, with the boundary conditions

$$\boldsymbol{v}^s = \textstyle\sum_{k,l=1}^{N} \bar{A}_k A_l \boldsymbol{\varphi}_{kl}^1 \text{ if } z = -1 \text{ or if } (x,y) \in \partial\Sigma, \tag{20}$$

$$\boldsymbol{v}^s\cdot\boldsymbol{e}_z = 0, \quad \partial\tilde{\boldsymbol{v}}^s/\partial z = \textstyle\sum_{k,l=1}^{N} \bar{A}_k A_l \boldsymbol{\varphi}_{kl}^2 \text{ at } z = 0. \tag{21}$$

Here and hereafter the *tildes* stand for the horizontal projection (of vectors or operators) and the vectors $\boldsymbol{\varphi}_{kl}^1$ and $\boldsymbol{\varphi}_{kl}^2$ are tangent to the unperturbed boundary and depend on the position along it; these vectors, omitted here in the interests of brevity, can be written as hermitian bilinear, differential operators on $(\boldsymbol{V}_k, \boldsymbol{V}_l)$, obtained from the analysis of the boundary layers mentioned above. In addition, for all k, l, m, n,

$$\delta_k = \mu_k C_g^{1/2}/\varepsilon, \quad d_k = (\omega - \Omega_k)/\varepsilon, \quad D = C_g/\varepsilon^2, \tag{22}$$

$$\alpha_{klmn} - \alpha_{klnm} = \alpha_{klmn} + \alpha_{lkmn} - \alpha_{mnlk} - \alpha_{nmlk} = 0, \tag{23}$$

$$\boldsymbol{g}_{kl} = \mathrm{i}\Omega^{-1}\boldsymbol{\nabla}\times(\bar{\boldsymbol{V}}_k \times \boldsymbol{V}_l). \tag{24}$$

Some remarks concerning the CASF equations above are now in order:

a. The amplitude equations (17) include terms accounting for inertia, viscous damping, detuning, cubic nonlinearity, parametric forcing and coupling with the streaming flow. Departure from the degenerate, multiple eigenvalue case is incorporated only in the detuning terms, a consistent procedure at leading order. Likewise, the viscous effects are only present explicitly in the damping term. Consequently, the cubic terms are conservative, and hence do not affect the energy of the system, given by (13), at leading order. The energy evolves according to

$$d\mathcal{E}/dT = -2\sum_{k=1}^{N} \delta_k |A_2|^2 + [\sum_{kl=1}^{N} \mathrm{i}\beta_{kl} \bar{A}_k \bar{A}_l + \text{c.c.}], \tag{25}$$

a result that follows on multiplying (17) by $\bar{A}_k$, adding the complex conjugate, adding up the results for $k = 1, \ldots, N$, and using (23)-(24). This equation also shows that the term accounting for the coupling with the streaming flow is also conservative *at leading order*, which is not surprising since the unscaled kinetic energy of the streaming flow is of order ε^2, while that of the surface waves is of order ε.

b. The new term accounting for the coupling with the streaming flow is of the same order as the usual cubic nonlinear terms and thus, except for some special cases in which this term vanishes, *it is inconsistent to neglect this term and retain only the cubic terms*, as frequently done in the literature. Nonetheless, if the system is initially at rest, the total streaming flow momentum (and hence the coupling term) can remain small for a long time, until the momentum diffuses (or is convected) from the boundary into the bulk; see remark *e* below.

c. In the generic case $N = 1$ the eigenfrequency Ω is algebraically simple, the phases of the eigenfunctions are constant, $\bar{\boldsymbol{V}}_1$ and $\boldsymbol{V}_1$ are colinear and (see (24)) $\boldsymbol{g}_{11} \equiv \boldsymbol{0}$. In this case the term accounting for the coupling with the streaming flow in the (one) amplitude equation (17) vanishes identically and *the evolution of A_1 decouples from the streaming flow*. Thus coupling requires the presence of at least two modes, as anticipated in §1.

d. The analysis above of the mean flow has been made for convenience in terms of the *Eulerian velocity*. This is just the average velocity on the fast timescale $t \sim 1$ and according to (11) is given by $\varepsilon(\boldsymbol{v}^s + \sum \bar{A}_k A_l \boldsymbol{\nabla} H_{kl})$ at leading order. The *mass transport*, or *Lagrangian*, velocity [6,15] is

$$\boldsymbol{v}^{mt} = \varepsilon(\boldsymbol{v}^s + \sum_{k,l=1}^{N} \bar{A}_k A_l \boldsymbol{\nabla} H_{kl}) + \boldsymbol{v}^{Sd} \tag{26}$$

at leading order, and is associated with the averaged trajectories of material elements; the difference between them is the Stokes drift, and is given by

$$\boldsymbol{v}^{Sd} = -\varepsilon \sum_{k,l=1}^{N} \bar{A}_k A_l \boldsymbol{g}_{kl} \tag{27}$$

at leading order, where $\boldsymbol{g}_{kl}$ is again given by (24). Note that both the inviscid mean flow and the Stokes drift are slaved to the surface waves, in contrast to the streaming flow. The mass transport velocity is relevant for comparison with flow visualizations (with an exposure time large compared to the forcing period) and, more generally, for analyzing the transport (and mixing) of passive scalars [16,17]. When the streaming flow is ignored this is usually because it is assumed that it is small compared to the Stokes drift [16]. However, except in some special cases (see below), this assumption is incorrect, as seen from (26)-(27). The normal component of the Eulerian mean-flow velocity at $z = 0$ is non-zero in general, but this does not lead to any mass transport across the unperturbed free surface. To see this we must consider the *mass transport velocity* (26). Eqs. (16) and (24) and standard formulas from vector analysis show readily that $\boldsymbol{\nabla} H_{kl} \cdot \boldsymbol{e}_z = \boldsymbol{g}_{kl} \cdot \boldsymbol{e}_z$ at $z = 0$, from which the result $\boldsymbol{v}^{mt} \cdot \boldsymbol{e}_z = 0$ at $z = 0$ follows at leading order on using (26)-(27).

e. The momentum equation (20) includes a non-standard term proportional to both the Lagrangian velocity and the Eulerian vorticity. This term is called *the vortex force* and does not force any flow by itself (it vanishes if $\boldsymbol{v}^s = \boldsymbol{0}$) but can enhance or inhibit the effect of the remaining forcing terms; in fact this term can destabilize some shear streaming flows produced by water waves, like those associated with Langmuir circulations in ocean and lakes, see [15] and references therein. Moreover, we have included a $O(\varepsilon)$ viscous term even though it is of higher order because it is required to satisfy the boundary conditions (20)-(21) that force the streaming flow. Since $\varepsilon \ll 1$, we must expect thin viscous boundary layers attached to the solid walls and the free surface, with thickness $\varepsilon^{1/2}$ and $\varepsilon^{1/3}$, respectively. According to (5) these are thicker than the boundary layers mentioned in §3 that lead to the boundary conditions (20)-(21), as required for consistency of the analysis. Note also that we must expect (at least) two timescales in the analysis of the CASF equations, namely a *convective timescale* $T \sim 1$ and a *viscous timescale*, $T \sim \varepsilon^{-1}$, which can lead to extremely slow dynamics.

f. According to (22a), the damping rate at leading order of the k-th mode is $\mu_k C_g^{1/2} > 0$, and results from viscous dissipation in the boundary layers mentioned above. Nevertheless, if the contact line is fixed, as we are assuming here, the coefficient μ_k is quite small for all but the first few low order modes [18] and (22a) must be replaced by $\delta_k = (\mu_k C_g^{1/2} + \mu_k' C_g)/\varepsilon$. If, in addition, C_g is not too small, then δ_k is in practice of order C_g/ε and $\delta_k \sim 1$ requires that $C_g \sim \varepsilon$. Then $\varepsilon D \sim 1$ and the two timescales mentioned above are no longer distinct.

4 Double-Mode Interaction in Almost-Square Containers

Let us assume now that the cross-section of the cylinder is a rectangle that is close to the square

$$\Sigma : |x| < L/2, \quad |y| < L/2, \tag{28}$$

that the inviscid eigenfrequency Ω is algebraically double, and that the associated eigenfunctions are $(\boldsymbol{V}_1, P_1, F_1)$ such that F_1 is even in the x-direction

and odd in the y-direction and $(\boldsymbol{V}_2, P_2, F_2)$, obtained from $(\boldsymbol{V}_1, P_1, F_1)$ by a reflection in the diagonal of the square. See [9,10] and [10,19,20,21,22,23] for experimental and theoretical analyses of this problem, all of them ignoring the streaming flow. Note that we are in the situation considered in §2-§3, with $N = 2$. After rescaling, the CASF equations (17)-(21) become

$$A'_{\pm}(\tau) = -[1 + \mathrm{i}(\Gamma \pm \Lambda)]A_{\pm} + \mathrm{i}(\alpha_1|A_{\pm}|^2 + \alpha_2|A_{\mp}|^2)A_{\pm} + \mathrm{i}\alpha_3\bar{A}_{\pm}A_{\mp}^2 + \mathrm{i}\Upsilon\bar{A}_{\pm} \pm \Omega\int_{-1}^{0}\int_{\Sigma}\boldsymbol{u}\cdot\boldsymbol{g}\,d\boldsymbol{x}A_{\mp}, \quad (29)$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \quad (30)$$

$$\partial\boldsymbol{u}/\partial\tau - [\boldsymbol{u} + \boldsymbol{H}(A_+, A_-) + \boldsymbol{G}(A_+, A_-)] \times (\boldsymbol{\nabla}\times\boldsymbol{u}) = -\boldsymbol{\nabla}p + Re^{-1}\Delta\boldsymbol{u}, \quad (31)$$

if $(x, y) \in \Sigma$ and $-1 < z < 0$, with boundary conditions

$$\boldsymbol{u} = (|A_+|^2 + |A_-|^2)\boldsymbol{\varphi}_1 + (|A_+|^2 - |A_-|^2)\boldsymbol{\varphi}_2 + (\bar{A}_+A_- + A_+\bar{A}_-)\boldsymbol{\varphi}_3 + \mathrm{i}(\bar{A}_+A_- - A_+\bar{A}_-)\boldsymbol{\varphi}_4 \text{ if either } z = -1 \text{ or } (x, y) \in \partial\Sigma, \quad (32)$$

$$\boldsymbol{u}\cdot\boldsymbol{e}_z = 0, \quad \partial\tilde{\boldsymbol{u}}/\partial z = \mathrm{i}(\bar{A}_+A_- - A_+\bar{A}_-)\boldsymbol{\varphi}_5 \quad \text{if } z = 0, \quad (33)$$

where $A_+ = A_1$, $A_- = A_2$, $\boldsymbol{u} = \boldsymbol{v}^s$, $\tau = T$, $\Gamma = \omega - (\Omega_1 + \Omega_2)/2$ and $\Lambda = (\Omega_1 - \Omega_2)/2$ up to $O(1)$ scaling factors, and

$$\boldsymbol{H} = \mathrm{i}(\bar{A}_+A_- - A_+\bar{A}_-)\boldsymbol{h}, \quad \boldsymbol{G} = \mathrm{i}(\bar{A}_+A_- - A_+\bar{A}_-)\boldsymbol{g}, \quad Re \sim \varepsilon^{-1} \gg 1. \quad (34)$$

Note that the detuning parameter Λ is a rescaled measure of the departure of the cylinder cross-section from the square (28). The presence of this parameter is the only manifestation of the departure from the square cross-section, the remaining parameters and functions being calculated for the square case. Thus if $\Lambda = 0$ the CASF equations are invariant under the actions

$$x \to -x, \quad A_+ \to -A_+, \quad \boldsymbol{u} \equiv (u_1, u_2, u_3) \to (-u_1, u_2, u_3), \quad (35)$$

$$x \leftrightarrow y, \quad A_+ \leftrightarrow A_-, \quad (u_1, u_2, u_3) \leftrightarrow (u_2, u_1, u_3), \quad (36)$$

which generate a group that is isomorphic to the symmetry group D_4 of a square. Invariance of (29)-(33) under this group requires that (i) $\boldsymbol{h}$, $\boldsymbol{g}$, $\boldsymbol{\varphi}_4$ and $\boldsymbol{\varphi}_5$ be invariant under the actions

$$x \to -x, \quad (u_1, u_2, u_3) \to (u_1, -u_2, -u_3), \quad (37)$$

$$x \leftrightarrow y, \quad (u_1, u_2, u_3) \to (-u_2, -u_1, -u_3), \quad (38)$$

(ii) $\boldsymbol{\varphi}_1$ be invariant under (35)-(36), (iii) $\boldsymbol{\varphi}_2$ be invariant under x- and y-reflections and under the action (37) and (iv) $\boldsymbol{\varphi}_3$ be invariant under reflections in the diagonals of the square and under the action (38).

4.1 Reflection Symmetries and Timescales; Qualitative Comparison with Experiments

Since $\boldsymbol{g}$ is invariant under (37)-(38), *the term describing the coupling with the streaming flow in (29) vanishes if the solution exhibits at least one of the reflection symmetries of the group* D_4 generated by (35)-(36) (at each value of τ if

the solution is time-dependent). Now we must take into account the symmetry properties of the container cross-section.

A. For squares ($\Lambda = 0$) there are three types of primary branches (all consisting of steady states), bifurcating from the flat state ($A_+ = A_- = 0$) at $\Upsilon = \sqrt{1+\Gamma^2}$: (i) pure states, of the form $A_+ \neq 0$, $A_- = 0$ and $A_+ = 0$, $A_- \neq 0$, (ii) standing mixed states of the form $A_+ = A_-$ and $A_+ = -A_-$ and (iii) progressive mixed states of the form $A_+ = \mathrm{i}A_-$ and $A_+ = -\mathrm{i}A_-$. The first two ones (the only ones encountered in [9]) are reflection-symmetric and thus do not involve the streaming flow. However, their stability properties do in general involve this flow, which could explain the quantitative differences encountered in [9] (see also [23]) between experiment and theory that ignored the streaming flow. The progressive mixed (steady) states exhibit an apparent rotation and have nonzero angular momentum, as explained in [10]; they do not possess any of the reflection symmetries in the group D_4, and thus the streaming flow (which also exhibits an apparent rotation, and has a nonzero angular momentum and vertical vorticity) affects not only their stability properties but also their shape. Note that these mixed modes are the ones for which the discrepancy between experimental observations and the theory for square cells by Feng and Sethna [10, Fig.5] is the largest.

B. For rectangles ($\Lambda \neq 0$) the *CASF* equations are no longer invariant under the group D_4, but only under the smaller group D_2 generated by (35) and

$$y \to -y, \quad A_- \to -A_-, \quad \boldsymbol{u} \equiv (u_1, u_2, u_3) \to (u_1, -u_2, u_3). \tag{39}$$

There are now only two primary branches of pure steady states, and the mixed steady states appear through secondary bifurcations. Fot this system Simonelli and Gollub [9] report seeing several types of interesting time-dependent states (some extremely slow, see below) that are not reflection-symmetric (at each τ) and thus should involve the streaming flow in an essential way. Thus far all attempts to explain these states using theory that neglects the presence of streaming flow have failed completely; the observations [9] constitute therefore a demanding test for any theory. Feng and Sethna's [10] theory compares well with the observations for the reflection-symmetric states (Fig. 6) but fails completely [10, Fig.6] for the steady progressive mixed states that should involve streaming flow. For example, where the theory predicted steady progressive mixed states Feng and Sethna observed in their experiments some extremely slow, amplitude modulated, rotating waves whose apparent chaotic nature could not be ascertained because of the extremely slow time scales involved.

In connection with the slow phenomena mentioned above, we note that the conclusion $Re \sim \varepsilon^{-1} \gg 1$ (see (34)) implies that the CASF equations (29)-(33) exhibit dynamics on two distinct timescales (§3, remark *e*), namely $\tau \sim 1$ and $\tau \sim \varepsilon^{-1} \gg 1$. If the observed dynamics occur on the latter timescale, the observations may be used to investigate the role of the streaming flow and of initial conditions, as we now discuss.

C. When the system is initially at rest, namely $\boldsymbol{u} \equiv \mathbf{0}$ at $\tau = 0$, a $O(1)$ tangential velocity will appear in a thin viscous boundary layer near the boundary where the vorticity is non-zero (due to the boundary conditions (32)-(33)),

but the total streaming flow momentum in the bulk will remain small for a long time, $\tau \sim Re^{-1}$, until it diffuses (and is convected) into the bulk from the boundary layers. During this transient the integral appearing in (29) remains small and the system approaches an attractor of the amplitude equations usually considered in the literature, namely those obtained when the streaming flow is ignored. But, after this long transient, the streaming flow comes into play and the solution converges to the true attractor of the full *CASF* system, frequently involving the streaming flow. Transients of this type have been reported in [9, §7.1] and in related laterally vibrated systems [24, Fig.8], and attractors reached on the viscous timescale might be involved in the experiment in [10]. If the viscous timescale is longer than the time span of the experiment and the initial streaming velocity is appropriately small, the effect of the streaming flow will not manifest itself during the experiment. This fact could explain why some of the observations by Feng and Sethna [10] agreed with a weakly nonlinear theory that ignored streaming flow, while those by Simonelli and Gollub [9] did not; the cross-section of the container in the former case was more than three times larger and the viscous timescale therefore at least 10 times longer (of the order of several hours).

D.- If the initial streaming flow velocity is nonzero and satisfies the boundary conditions, the initial vorticity is also nonzero, and the streaming flow affects the dynamics of the surface waves from the very beginning. During an initial long transient on the convective timescale viscous diffusion in the momentum equation (31) can be ignored; the resulting system of simplified *CASF* equations can exhibit "attractors" that need not be close to the true attractors of the system, which will be reached only on the viscous timescale. Transients of this type might be responsible for the striking behavior reported in [9, Fig.16] and described as 'structural instability', namely, that in these long transients the behavior of the system depends strongly on the initial condition for the streaming flow velocity, which is not controlled in the experiments [9].

This work was supported by NASA (Grant NAG3-2152) and DGES (Grant PB97-0556).

References

1. M. Faraday, On the forms and states assumed by fluids in contact with vibrating elastic surfaces Phil. Trans. Roy. Soc. London **121** (1831), 319–340.
2. J. Miles and D. Henderson, Parametrically forced surface waves, Annu. Rev. Fluid Mech. **22** (1990), 143–165.
3. M.C. Cross and P.C. Hohenberg, Pattern formation outside of equilibrium Rev. Modern Phys. **65** (1993), 851-1112.
4. S. Fauve, Parametric instabilities, in *Dynamics of Nonlinear and Disordered Systems*, G. Martínez-Mekler and T.H. Seligman eds., World Scientific, pp.67-115, 1995.
5. H. Schlichting, Berechnung ebener periodischer Grenzschichtströmungen, Phys. Z. **33** (1932), 327-335.

6. M.S. Longuet-Higgins, Mass transport in water waves, Phil. Trans. Roy. Soc. A **245** (1953), 535-581.
7. M. Higuera, J.A. Nicolás and J.M. Vega, Coupled amplitude-streaming flow equations for the evolution of counter-rotating, nearly-inviscid surface waves in finite axisymmetric geometries. Preprint, 2000.
8. F. Mashayek and N. Ashgriz, Nonlinear oscillations of drops with internal circulation, Phys. Fluids **10** (1998), 1071-1082.
   J. Fluid Mech. **158**(1985)381-398.
9. F. Simonelli and J.P. Gollub, Surface wave mode interactions: effects of symmetry and degeneracy, J. Fluid Mech. **199** (1989), 471-494.
10. Z.C. Feng and P.R. Sethna, Symmetry breaking bifurcations in resonant surface waves, J. Fluid Mech. **199** (1989), 495-518.
11. G.W. Young and S.H. Davis, A plate oscillating across a liquid interface: effects of contact-angle hystheresis, J. Fluid Mech. **174** (1987), 327-356.
12. J. Bechhoefer, V. Ego, S. Manneville and B. Johnson, An experimental study of the onset of parametrically pumped surface waves in viscous fluids, J. Fluid Mech. **288** (1995), 325-350.
13. H. Lamb, *Hydrodynamics*, Cambridge University Press, 1932.
14. J.A. Nicolás and J.M. Vega, Weakly nonlinear oscillations of axisymmetric liquid bridges, J. Fluid Mech. **328** (1996), 95-128.
15. A.D.D. Craik, *Wave Interactions and Fluid Flows*, Cambridge Univ. Press, 1985.
16. M. Umeki, Particle transport by angular momentum on three-dimensional standing surface waves, Phys. Rev. Lett. **67** (1991), 2650-2653.
17. Z.C. Feng and S. Wiggins, Fluid particle dynamics and Stokes drift in gravity and capillary waves generated by the Faraday instability, Nonlinear Dynamics **8** (1995), 141-160.
18. C. Martel, J.A. Nicolás and J.M. Vega, Surface-wave damping in a brimful circular cylinder, J. Fluid Mech. **360** (1998), 213-228. See also Corrigendum, **373** (1998), 379.
19. M. Silber and E. Knobloch, Parametrically excited waves in square geometry, Phys. Lett. A **137** (1989), 471-494.
20. M. Nagata, Nonlinear Faraday resonance in a box with a square base, J. Fluid Mech. **209** (1989), 265-284.
21. M. Nagata, Chaotic behavior of parametrically excited surface waves in square geometry, Eur. J. Mech. B/Fluids **10** (1991), 61-66.
22. M. Umeki, Faraday resonance in rectangular geometry, J. Fluid Mech. **227** (1991), 161-192.
23. D. Armbruster, J. Guckenheimer and S. Kim, Resonant surface waves in a square container, in *Differential Equations and Computer Algebra*, M. Singer ed., Academic Press, pp.61-76, 1991.
24. M. Funakoshi and S. Inoue, Surface waves due to resonant horizontal oscillation, J. Fluid Mech. **192** (1988), 219-247.

Stationary and Oscillatory Flow of a Charged Liquid Around a Blade Electrode

Francisco J. Higuera

E.T.S. Ingenieros Aeronáuticos, Pza. Cardenal Cisneros 3, 28040 Madrid, Spain

1 Introduction

The presence of space charge in a dielectric liquid subject to an electric field leads to a Coulomb force that may set the liquid into motion. Such motion is of interest in a number of applications including electrostatic precipitators, electrohydrodynamic (EHD) ion-drag pumps, and several instances of EHD laminar and turbulent mixing, which depend on the transport of charge, momentum or heat across the liquid [1,2].

Only ions, not electrons, contribute to the electric current. Field-enhanced dissociation of electrolytic impurities [3,4,5] and injection of ions by electrochemical reactions at metallic electrodes [6,7,8] are the most important mechanisms of generation of charge in a wide range of electric fields. The second of these tends to dominate at large fields [9,10,11] and then, typically, the rate-limiting process for injection is the escape of ions over the image-force barrier at the metal surface. This leads to an injection current or charge density that increases very rapidly with the local electric field.

Needle-plane and knife-plane electrode configurations have been used to generate high current densities in nonpolar liquids, because they feature locally high electric fields that enhance charge generation and motion. The steady current-voltage characteristics of these electrode sets display ohmic and quasi-ohmic regimes followed by a rapid increase of the current with increasing voltage [11,12]. A current pulse regime also exists for moderately rounded electrode tips and values of the electric field above a certain threshold. When the peaky electrode is negative, this regime is regular and analogous to Trichel oscillations in gases, which led Denat et al. [13] and Henández-Ávila et al. [14] to postulate that processes analogous to electron avalanches, field emission enhancement, and quenching by negative space charge, which are well-known in gases, might be occurring also in liquid phase.

Here, building on a previous analysis [15], the flow and charge distribution around a blade-shaped electrode will be discussed assuming that the only mechanism of generation of charge is injection by the electrode, in such a manner that the charge density at each point of the electrode surface is a known, rapidly increasing function of the electric field. The stationary flow and charge distribution predicted by this model in the absence of fluid inertia have been determined elsewhere [16], showing that convection dominates the transport of charge far away from the electrode and confines the space charge to a thin streak, and that

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 338–348, 2001.
© Springer-Verlag Berlin Heidelberg 2001

the electric field induced by this streak and its image by the electrode dominates eventually the flow and limits the electric current. This analysis has been extended in [15] to take into account the inertia of the liquid and the formation of an EHD plume. Here the inertia will be neglected and attention will be confined to the case of small currents (as defined in the following section), in which the harmonic field, due to the voltage applied between the injecting electrode and another collecting electrode, is large compared with the field induced by the space charge. In this case the space charge-induced field may still influence injection, owing to the high sensitivity of the injection law, but its influence comes about as a collective effect of the charge in the whole interelectrode gap rather than of the charge in the vicinity of the injecting electrode. A delay, determined by the liquid velocity, is thereby introduced between injection of charge and the field this charge creates on the electrode, which may lead to oscillations in the framework of the present simple injection model. Order of magnitude estimates of the frequency and other characteristics of these oscillations are worked out and backed by numerical computations using a simplified condition that simulates the effect of the collecting electrode without actually introducing geometrical details of the cell.

2 Formulation

Consider the two-dimensional flow induced by Coulomb's force in an electrically charged dielectric liquid between two metallic electrodes subject to a constant voltage. One of the electrodes is a blade of curvature radius r_0 at its tip, and injects charge into the liquid according to a law $q = f(E)$ that gives the charge density at each point of the electrode surface as a rapidly increasing function of the magnitude of the local electric field $E = |\boldsymbol{E}|$. The electric charge is then convected by the liquid and drifts relative to the liquid with a velocity proportional to the electric field, until it reaches the other, collecting electrode. The ratio of the interelectrode distance to twice r_0 will be denoted by R and will be assumed to be large in what follows. Advancing the fact that the region of characteristic size r_0 around the tip of the injecting electrode (where the electric field is highest) plays an essential part in the solution for $R \gg 1$, the characteristic values of the different magnitudes in this region will be determined and used to nondimensionalize the problem.

Let E_c and q_c denote the orders of magnitude of the electric field and the charge density at distances of $O(r_0)$ from the tip of the electrode. The condition for the space charge in this region to contribute to the local field is $q_c = O(\epsilon E_c/r_0)$ (from Poisson's law $\boldsymbol{\nabla} \cdot \boldsymbol{E} = q/\epsilon$, where ϵ is the permittivity of the liquid). Carrying this estimate of the charge density into the injection law gives $\epsilon E_c/2r_0 = f(E_c)$ in orders of magnitude, where the factor 2 is introduced for convenience; this will be taken to define E_c hereafter. The characteristic velocity v_c of the liquid around the electrode can be obtained from the balance of Coulomb force, of $O(q_c E_c)$, and viscous force, of $O(\mu v_c/r_0^2)$, where μ is the viscosity of the liquid. This balance gives $v_c = \epsilon E_c^2 r_0/\mu$.

In what follows, distances will be scaled with $2r_0$, time with $\mu/\epsilon E_c^2$, and the variables (ψ, ω, ϕ, q), where ψ is the stream function, ω is the vorticity, and ϕ is the electric potential, so that $\boldsymbol{E} = \boldsymbol{\nabla}\phi$, with the factors ($4\epsilon E_c^2 r_0^2/\mu$, $\epsilon E_c^2/\mu$, $E_c r_0$, $\epsilon E_c/2r_0$), respectively. The definition of E_c given above implies $f(1) = 1$ in nondimensional variables. The large sensitivity of the nondimensional injection law ($\mathrm{d}f/\mathrm{d}E \gg 1$) implies that the electric field cannot be much larger than unity anywhere on the electrode, because q would be locally very high, screening the electrode and reducing E, and that q is very nearly zero in the non-injecting regions where E falls appreciably below unity. Linearizing $\ln f$ about $E = 1$ gives the simplified injection law $q = \exp\{\gamma(E-1)\}$, which contains the single parameter $\gamma = \mathrm{d}\ln f/\mathrm{d}E|_{E=1} \gg 1$. This law displays the features mentioned above and will be used in what follows. In the asymptotic limit $\gamma \to \infty$ it reduces to ($E = 1$, $q > 0$) in the injecting region of the electrode and to ($E < 1$, $q = 0$) in the rest of the electrode surface.

Leaving out the inertia of the fluid, which is justified when the Reynolds number $\mathrm{Re} = v_c r_0/\mu = 2\rho\epsilon E_c^2 r_0^2/\mu^2$ is small, the nondimensional equations governing the flow around the electrode tip are

$$\nabla^2\psi = -\omega, \tag{1}$$

$$\frac{\partial q}{\partial x}\frac{\partial \phi}{\partial y} - \frac{\partial q}{\partial y}\frac{\partial \phi}{\partial x} + \nabla^2\omega = 0, \tag{2}$$

$$\nabla^2\phi = q, \tag{3}$$

$$\frac{\partial q}{\partial t} + \boldsymbol{\nabla}\cdot[q(\boldsymbol{v} + \kappa\boldsymbol{E})] = 0, \tag{4}$$

where $\kappa = \mu K/2\epsilon E_c r_0$, K being the charge mobility. Here the electrode is assumed to be a parabola; x and y are nondimensional distances measured from the focus of the parabola along its axis and normal thereto (see figure 1 below), so that $x = 1 - y^2/4$ on the electrode; and $\boldsymbol{v} = (u, v) = (\partial\psi/\partial y, -\partial\psi/\partial x)$. Later, polar coordinates $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan(-y/x)$, with $\theta = \pi$ at $x > 0$, $y = 0$, will also be used. Diffusion of charge is neglected.

The boundary conditions at the injecting electrode are

$$x = 1 - \frac{y^2}{4}: \quad \psi = |\boldsymbol{\nabla}\psi| = \phi = 0, \qquad q = f(|\boldsymbol{\nabla}\phi|) = \exp\{\gamma(|\boldsymbol{\nabla}\phi| - 1)\} \tag{5}$$

Boundary conditions are also needed at the collecting electrode ($x = 1 + R$ in a blade-plane configuration) and at the side walls of the cell confining the liquid. The conditions at the collecting electrode are that the velocity of the liquid is zero and that $\phi = V$, the nondimensional voltage applied between the two electrodes. The conditions at the side walls depend on the electrical properties of these walls and the geometrical configuration of the cell, both of which are thereby brought into the formulation of the problem.

The nondimensional electric current is $J = \int(u + \kappa\partial\phi/\partial x)q\,\mathrm{d}y$, where the integral is taken across the whole cell. Under stationary conditions this quantity does not depend on x [from (4)] and is equal to $\kappa\int Ef(E)\,\mathrm{d}s$ [from (5)], where the integral is taken over the surface of the injecting electrode.

The sources of the electric field in the region $r = O(1)$ are the voltage applied between the electrodes and the space charge in this region and elsewhere in the cell. In the absence of space charge, the electric potential would be $\Phi_\infty \phi_h$ at leading order for $r \ll R$, where $\phi_h = \sqrt{(r+x)/2} - 1$ solves $\nabla^2 \phi_h = 0$ (the subscript h stands for harmonic) and is constant on parabolas confocal with the electrode, and $\Phi_\infty = a\left(V/R^{1/2}\right)\left[1 + O(R^{-1/2})\right]$, with the proportionality constant $a = O(1)$ dependent on the shape of the cell. If the space charge in the bulk of the cell, appearing on the right hand side of (3), is important, then the harmonic function $\Phi_\infty \phi_h$ is only a part of the potential, and the value of Φ_∞ seen by the region of $r = O(1)$ may be much smaller than $V/R^{1/2}$.

The solution in this region, which is the main subject of the present work, is determined by (1)–(5) plus certain asymptotic conditions for $r \to \infty$, obtained by matching the inner region $r = O(1)$ and the bulk of the cell. Such conditions were given in [16] for $J = O(1)$ and will be worked out in Sec. 3 for $J \ll 1$. It is worth advancing here that the matching conditions turn out to be largely independent of the configuration of the cell and express that, insofar as the inner region is concerned, the motion of the liquid confines the space charge to a thin streak and that the electric field for $r \gg 1$ (but of course $\ll R$) is dominated by the contribution of this charged streak and its image in the injecting electrode, if $J = O(1)$, or by the applied voltage and the far charge, if $J \ll 1$. Figure 1 shows a sample numerical solution in the inner region.

The stationary form of (1)–(5), supplemented with matching conditions for $r \to \infty$, should determine the stationary solution in the inner region, giving in particular the current J as a function of Φ_∞ and the parameters κ and γ. The solution of this problem has been computed elsewhere[16] and extended in [15] to cases with Re > 0. The asymptotic limit $J \to 0$ will be analyzed in Sec. 3 and oscillatory regimes existing for $J \ll 1$ will be discussed in Sec. 4.

3 Stationary States

Small values of J are associated with small values of the space charge, so that the electric field induced by this charge should be expected to be small compared with the harmonic field $\boldsymbol{E}_h = \Phi_\infty \boldsymbol{\nabla} \phi_h$ in the inner region $r = O(1)$ and up to some large distance from the electrode (to be determined below), where the space charge is already confined to a thin streak around the symmetry plane.

The field at the surface of the injecting electrode is then $E \approx E_h = \Phi_\infty |\boldsymbol{\nabla} \phi_h| = (\Phi_\infty/2)\left[1 - y^2/8 + O(y^4)\right]$ around the tip, leading to a charge density $q = f(E_h) = q_0 \exp\left(-\gamma \Phi_\infty y^2/16\right)\left[1 + O(\gamma \Phi_\infty y^4)\right]$, where $q_0 = \exp\{\gamma(\Phi_\infty/2 - 1)\}$ is the charge density at the tip. The values of Φ_∞ of interest are close to 2 when $\gamma \gg 1$, because otherwise q_0 would be either exponentially small, leading to a negligible current, or exponentially large, invalidating the assumption that the harmonic field dominates. In addition, in the case of $\gamma \gg 1$, noticeable injection occurs only in the region of the electrode surface around the tip where

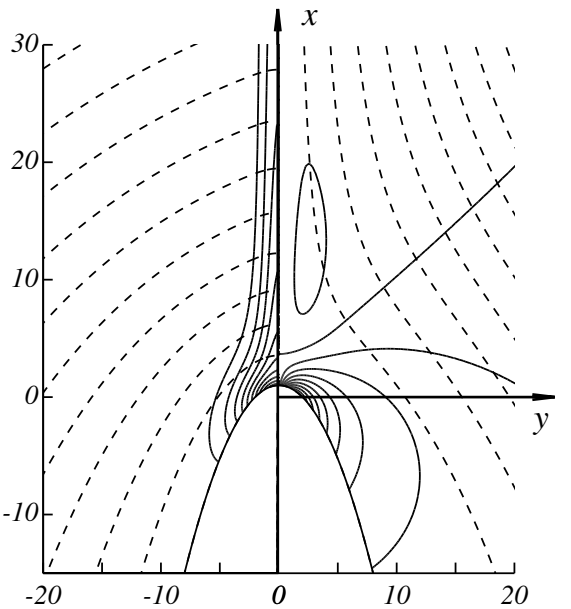

Fig. 1. Numerical solution for $\Phi_\infty = 2$, $\kappa = 0.5$ and $\gamma = 20$. Left: Ten contours of q between 0.05 and 0.5 (solid) and ten contours of ϕ between 2.5 and 25 (dashed). Right: Eleven contours of ω between -0.12 and 1.08 (solid) and eight contours of ψ between 10 and 80 (dashed).

$y = O(\gamma^{-1/2})$, so that the current is

$$J = \int_{-\infty}^{\infty} \kappa E q \, \mathrm{d}y \approx 2^{3/2} \pi^{1/2} \frac{\kappa q_0}{\gamma^{1/2}}, \tag{6}$$

where Φ_∞ has been set equal to 2 everywhere except in q_0.

The electric field induced at the electrode by the space charge in the region $r = O(1)$ is of order $q_0/\gamma^{1/2}$ [from (3)]. This field will affect injection when it becomes of order γ^{-1}, which happens for $q_0 \sim \gamma^{-1/2}$. However, the first influence of space charge on injection comes from the charge in the bulk of the cell, not in the inner region $r = O(1)$, as will be seen next.

Consider the flow for $x \gg 1$ but still very small compared with R. The electric field acting on the charged streak is $E_h = \Phi_\infty / 2x^{1/2}$, pointing in the x-direction. Calling $u_0(x)$ the velocity of the liquid at the symmetry plane and advancing that $u_0 \gg \kappa E_h$ for $x \gg 1$ (cf. (7) below), the current can be written as $J \approx \int_{-\infty}^{\infty} uq \, \mathrm{d}y \approx u_0 \int_{-\infty}^{\infty} q \, \mathrm{d}y$, and the Coulomb force acting on the charged streak is $E_h \int_{-\infty}^{\infty} q \, \mathrm{d}y = E_h J / u_0$. In the absence of inertia, this Coulomb force must be balanced by the viscous force acting on the sides of the streak, of order u_0/x, whence $u_0 = O(\Phi_\infty^{1/2} J^{1/2} x^{1/4})$. To make these estimates more precise, notice that $\nabla^4 \psi = 0$ outside the streak, and the solution of this equation that is proportional to $r^{5/4}$ and satisfies the boundary conditions $\psi = \partial\psi/\partial\theta = 0$ at

the electrode surface ($\theta \approx 0$ for $r \gg 1$ in polar coordinates) and $\psi = 0$ at the symmetry plane $\theta = \pi$ is $\psi = Ar^{5/4}\left(\cos\frac{3}{4}\theta - \cos\frac{5}{4}\theta\right)$, with A arbitrary. Using this solution, $u_0 = -r^{-1}\partial\psi/\partial\theta|_{\theta=\pi} = 2^{1/2}Ax^{1/4}$ and the viscous force on the two sides of the streak is $2r^{-2}\partial^2\psi/\partial\theta^2|_{\theta=\pi} = 2^{1/2}A/x^{3/4}$, so that the previous balance gives $A = (\Phi_\infty J)^{1/2}/2$.

The electric field induced by the charged streak and its image in the electrode can now be computed. The space charge density integrated across the streak is $Q(x) = \int_{-\infty}^{\infty} q\,\mathrm{d}y = J/u_0 = 2^{1/2}(J/\Phi_\infty)^{1/2}/x^{1/4}$ for $x \gg 1$. Let $Q_i(x)$ denote the image charge in the injecting electrode for $(-x) \gg 1$, and notice that the image charge in the collecting electrode, as well as the successive images in both electrodes, lead to contributions to the electric field which are small, of $O\left[(x/R)^{1/4}\right]$, relative to the contributions of $Q(x)$ and $Q_i(x)$ in the region of concern here ($1 \ll x \ll R$), which is still the vicinity of the injecting electrode. Since this electrode can be replaced by the plane $y = 0$ when $(-x) \gg 1$, the condition that the component of the electric field tangent to the electrode be zero is

$$\int_0^\infty \frac{Q(x')}{x - x'}\,\mathrm{d}x' + \int_{-\infty}^0 \frac{Q_i(x')}{x - x'}\,\mathrm{d}x' = 0$$

at large distances from the tip (here the second integral is a principal value), which is satisfied for $Q_i(x) = 2^{1/2}Q(-x)$. (This can be verified by substituting $Q(x) = constant/x^{1/4}$ and $Q_i(x) = 2^{1/2}constant/(-x)^{1/4}$ in the integrals above, making the changes of variables $x' = -x\,\xi$ in the first integral and $x' = x\,\xi$ in the second, and evaluating analytically the resulting integrals.) The field induced on the streak far away from the electrode tip is then

$$E_{sc} = \frac{1}{2\pi}\int_0^\infty \frac{Q(x')}{x - x'}\,\mathrm{d}x' + \frac{1}{2\pi}\int_{-\infty}^0 \frac{Q_i(x')}{x - x'}\,\mathrm{d}x' = \frac{(J/\Phi_\infty)^{1/2}}{2^{1/2}x^{1/4}},$$

where the first integral is now a principal value and the subscript sc is used hereafter to denote the space charge induced field and its potential. Outside the streak $\boldsymbol{E}_{sc} = \boldsymbol{\nabla}\phi_{sc}$ with $\nabla^2\phi_{sc} = 0$ and $\phi_{sc} = 0$ at the electrode, which, along with the previous expression of E_{sc} at the streak, determines

$$\phi_{sc} = \frac{4}{3}\left(\frac{J}{\Phi_\infty}\right)^{1/2} r^{3/4}\sin\frac{3}{4}\theta.$$

Summarizing, the asymptotic conditions for $r \to \infty$ are, in the present case,

$$\left.\begin{array}{l} \psi \sim \frac{1}{2}(\Phi_\infty J)^{1/2} r^{5/4}\left(\cos\frac{3}{4}\theta - \cos\frac{5}{4}\theta\right) + \dots \\ \omega \sim -\frac{1}{2}(\Phi_\infty J)^{1/2} r^{-3/4}\cos\frac{3}{4}\theta + \cdots \\ \phi - \Phi_\infty\phi_h \sim \frac{4}{3}\Phi_\infty^{-1/2} J^{1/2} r^{3/4}\sin\frac{3}{4}\theta + \cdots \\ q = 0 \quad \text{outside the streak} \end{array}\right\} \tag{7}$$

As can be seen, the space charge contribution to the electric potential increases with x or r more rapidly than the harmonic potential [$\phi_h = O(x^{1/2})$]. The

two contributions become of the same order for $x = O(J^{-2})$ [when $\Phi_\infty = O(1)$], and the charge-induced field dominates further away, which invalidates the analysis leading to (7) in that far region. The crossover distance is larger than the size R of the cell when $J \ll R^{-1/2}$, in which case the harmonic field dominates throughout the cell and the electric potential at the collecting electrode, equal to the applied voltage V, is

$$V = aR^{1/2}\Phi_\infty + bR^{3/4}J^{1/2}, \tag{8}$$

where a and b are order unity constants dependent of the shape of the cell. Here $\Phi_\infty = 2$ has been used in the second term of (8), which is due to the space charge. Leaving out this term, the value of Φ_∞ would be entirely determined by the applied voltage, which should be close to $2aR^{1/2}$ in the conditions under scrutiny. However, for a constant V, the small reduction of Φ_∞ due to the space charge, equal to $-bJ^{1/2}R^{1/4}/a$, may be enough to affect injection if it is of $O(1/\gamma)$ or larger; i.e. if $J = O(\gamma^{-2}/R^{1/2})$ or larger, or $q_0 = O(R^{-1/2}\gamma^{-3/2})$ or larger. The J–V characteristic, obtained eliminating Φ_∞ between (6) and (8), is

$$J^2 = 8\pi\frac{\kappa^2}{\gamma}\exp\left(\frac{\delta V}{aR^{1/2}} - \frac{b\gamma R^{1/4}}{a}J^{1/2}\right) \quad \text{for} \quad J \ll \min(R^{-1/2}, \kappa/\gamma), \tag{9}$$

where $\delta V = \gamma(V - 2aR^{1/2})$.

4 Oscillatory States

Equations (1)–(3) imply that the electric field and the velocity of the liquid everywhere in the cell, and in particular at the electrode, adjust instantly to the space charge distribution. This distribution, however, needs not be stationary and smooth, as assumed above. Instead, the charged streak lying between the electrodes could develop alternating accumulations and defects of charge that would induce a similar modulation of the electric potential. When one of these lumps of charge approaches the collecting electrode, fixed at a constant voltage $\phi = V$, a variation of Φ_∞ results, which may fire a new lump at the injecting electrode and thus sustain the oscillations.

Let Q_l be the charge of a lump, a negative value meaning a defect of charge relative to the mean distribution. In two space dimensions, the electric potential due to a lump approaching the collecting electrode is of order Q_l, up to logarithms. Adding an extra term of $O(Q_l)$ to the right hand side of (8), the variation of Φ_∞ necessary to keep V constant would be of order $Q_l/R^{1/2}$. On the other hand, this variation should be of order γ^{-1} in order to influence significantly q_0 and the injected current, so that $Q_l = O(\gamma^{-1}R^{1/2})$. With a mean current J, the duration of a pulse containing this charge would be $T = O(\gamma^{-1}R^{1/2}/J)$, and the length x_l of the lump, which is convected by the liquid with a velocity of order $u(x_l) = O(J^{1/2}x_l^{1/4})$ [from (7)] while being injected, satisfies $x_l \sim u(x_l)T$, whence $x_l = O(\gamma^{-4/3}R^{2/3}/J^{2/3})$. Later, when the lump is in

the bulk of the cell, its velocity is of order $u(R) = O(J^{1/2}R^{1/4})$ (extrapolating (7) to distances of order R), so that the separation between adjacent lumps is $\Delta x_l = O[u(R)T] = O(\gamma^{-1}R^{3/4}/J^{1/2})$, and the number of lumps simultaneously present in the cell is of order $R/\Delta x_l = O(\gamma R^{1/4}J^{1/2})$. This number should be of order unity or large, which requires a current of order $\gamma^{-2}/R^{1/2}$ or larger. Thus the estimate of the minimum current in the oscillatory regime coincides with the one obtained in the last paragraph of Sec. 3 for the space charge in the bulk to first affect stationary injection. The coincidence is not surprising since there is no qualitative difference, as far as these estimations are concerned, between a stationary state and a oscillatory state with only one or a few lumps of charge in the cell.

These qualitative estimates agree with some of the experimental results of Denat et al. [13] for cyclohexane in point-plane electrode geometry at pressures high enough to prevent the formation of bubbles during the bursts of current. Specifically, these authors find that the repetition frequency of the bursts $(1/T)$ is proportional to the mean current J and that the charge per burst (Q_l) is independent of the applied voltage. They also observe that the burst period T is only a small fraction of the transit time of the charge between the two electrodes, which is in line with the presence of many bursts in the cell. At variance with the present estimates, the experiments of Denat et al. show that the electrode gap, which these authors varied in the range 0.5–3 mm, has no influence on the pulse regime, whereas the point radius, which was between 0.5 μm and 10 μm in the pulse regime, is the parameter with the strongest influence. The ratio of these two parameters is analogous to R here. The discrepancy may be due to differences between the planar geometry considered here and the axisymmetric geometry of Denat et al. experiments. In fact, the factor $R^{1/2}$ in the harmonic potential contribution to (8) and in the estimates of Q_l and T would become $\ln R$ for an axisymmetric paraboloidal electrode.

A necessary condition for the oscillations just described to exist is that the lumps of charge preserve their identity during their journey across the cell. Seen as successive accumulations and defects of charge of magnitude $\pm Q_l$ relative to a mean distribution, each accumulation induces on an adjacent defect, a distance Δx_l apart, an electric field of order $E_l = Q_l/\Delta x_l$ (in two dimensions). This field, in turn, gives an attractive force of order $E_l Q_l/(\Delta x_l)^2$ per unit area, and the balance of this force and the viscous force $v_l/(\Delta x_l)^2$, where v_l is the order of the relative velocity due to this effect, requires $v_l = O(E_l Q_l)$. Thus, the time it takes to smear out a lump is $t_s = O(\Delta x_l/v_l) = O(R^{1/2}/J)$, which is to be compared with the transit time across the cell $t_t = O[R/u(R)] = O(R^{3/4}/J^{1/2})$. The condition $t_s \gg t_t$ amounts to $J \ll 1/R^{1/2}$, which is an upper bound on the current for the present oscillations to exist. This upper bound coincides with the estimate in the previous subsection of the current above which the field of the stationary space charge overcomes the harmonic field in the bulk of the cell. When $1/R^{1/2} \ll J \ll 1$ the harmonic field dominates only in a reduced region, which could perhaps host charge lumps and oscillations if the rest of the cell acts as a distributed electrode. Such oscillations, if they exist, would be

independent of R, in line with the experimental observations of Denat et al. [13]. This possibility will not be discussed here.

The condition $\phi = V$ at the collecting electrode plays an important role in the oscillatory regime. In the remainder of this section this boundary condition will be replaced by a modeled one, in order to check the existence of oscillations without having to deal with the details of any particular cell, and problem (1)–(5) will be simplified taking advantage of the small value of J.

It was mentioned before that the oscillatory variation of Φ_∞ seen by the inner region is of the order of $R^{-1/2}$ times the electric potential due to a lump of charge, which in turn is of the order of Q_l at distances from the lump of the order of the characteristic lump spacing Δx_l. This electric potential can be gauged by computing the space average of the space charge-induced potential $\phi_{sc} = \phi - \Phi_\infty \phi_h$ over a large surface (a contour in two dimensions) surrounding the injecting electrode. A lump of charge crossing this surface, or at a distance of $O(\Delta x_l)$ from it, will lead to a contribution of order $Q_l \Delta x_l / R_\infty$ to the average of ϕ_{sc}, where R_∞ is the characteristic size of the averaging surface. The model boundary condition consists in the assumption that the variation of Φ_∞ about its stationary value is proportional to the variation of the average of ϕ_{sc} about its stationary value.

Consider a reference stationary state with a current $J_e \ll 1$, to which (6) associates a value $\Phi_{\infty e} = 2 + (2/\gamma) \ln \left[(J_e \gamma^{1/2}) / (2^{3/2} \pi^{1/2} \kappa) \right]$. Rescale the variables ψ, ω and $\phi_{sc} = \phi - \Phi_\infty \phi_h$ and the parameter κ with $J_e^{1/2}$, q with $J_e^{1/2} \gamma^{1/2}$, and the time with $J_e^{-1/2}$. In the asymptotic limit $J_e \to 0$, the rescaled variables are of order unity for $r = O(1)$ and satisfy (1)–(4) with the electric field replaced by $2 \boldsymbol{\nabla} \phi_h$ in (2) and (4). The first three boundary conditions (5) do not change in this limit, while the rescaled charge density at the injecting electrode is given by

$$q = q_e \exp(\gamma \delta \Phi_\infty / 2), \tag{10}$$

where $q_e = \exp(-\gamma y^2/8) / (2^{3/2} \pi^{1/2} \kappa)$ is the rescaled stationary charge density (κ being here the rescaled mobility) and $\delta \Phi_\infty$ is the variation of Φ_∞ about its stationary value $\Phi_{\infty e}$.

The model boundary condition is

$$\delta \Phi_\infty = -\alpha \left\langle \phi_{sc} - \phi_{sc_e} \right\rangle, \tag{11}$$

where ϕ_{sc_e} is the stationary ϕ_{sc} and the brackets denote the space average over a contour far away from the injecting electrode, which will be taken here as $x = R_\infty - y^2/R_\infty$ with R_∞ large. Finally, α in (11) is a free parameter of the model, which should be of order $(R_\infty/R)(J_e/\gamma^{-2} R^{-1/2})/(\gamma R^{3/4}) \ll 1$ [because $\delta \Phi_\infty = O(\gamma^{-1})$ and $\langle \phi_{sc} - \phi_{sc_e} \rangle = O(Q_l \Delta x_l / R_\infty J_e^{1/2}) = O\left[(R/R_\infty)(R^{1/4}/\gamma^2 J_e) \right]$ in rescaled variables, using the estimates of Q_l and Δx_l above]. Thus α gauges both the stationary current and the size of the modeled cell.

The rescaled equations (1)–(4) are solved subject to these conditions and to the far field conditions (7) with J (which is rescaled with J_e) evaluated as a time average and $\Phi_\infty = 2$ on the right hand sides. These latter conditions are imposed at the contour used to evaluate $\delta \Phi_\infty$.

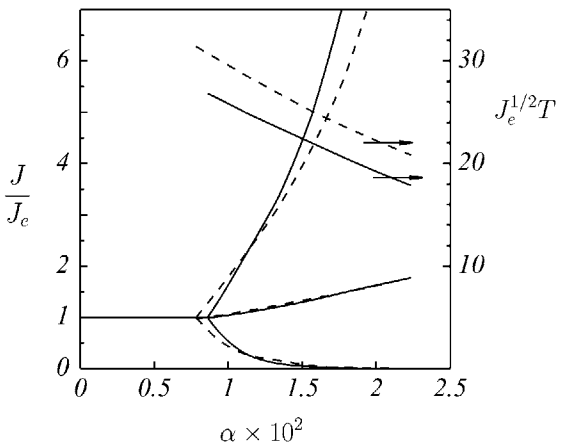

Fig. 2. Minimum, maximum and mean currents as functions of α for $\gamma = 20$, $J_e = 5 \times 10^{-2}$ and $R_\infty = 49$. Solid curves are for $\kappa/J_e^{1/2} = 4.47$ and dashed curves for $\kappa/J_e^{1/2} = 0.89$. The rescaled period of the oscillation is given on the right ordinate.

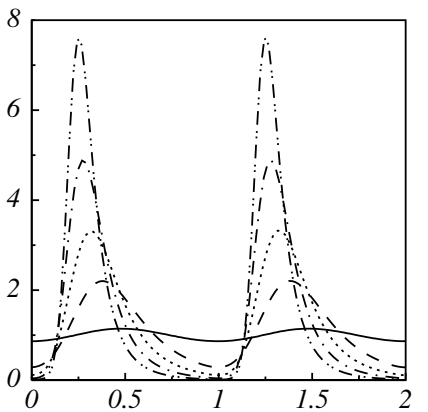

Fig. 3. Injected current during two cycles of the oscillation as a function of time, scaled with the period, for $\alpha = 0.0089$ (solid), 0.0112 (dashed), 0.0134 (dotted), 0.0157 (dash-and-dot), and 0.0179 (dash-double dot). Values of other parameters are $\kappa/J_e^{1/2} = 4.47$, $\gamma = 20$, $J_e = 5 \times 10^{-2}$, and $R_\infty = 49$.

The stationary state for a given J_e was first computed numerically by iterating on Φ_∞ until the rescaled injected current is equal to unity. Then an arbitrary perturbation was added to the solution and its evolution was computed for different values of α. These computations show that the perturbation decays when α is smaller than a certain critical value, dependent on κ and γ, and becomes time periodic when α is larger than this critical value. The minimum, maximum

and mean currents in the oscillatory regime are plotted in figure 2 as functions of α for $\gamma = 20$ and two different values of the rescaled mobility. In these computations $R_\infty = 49$ and the stationary states were computed with $J_e = 5 \times 10^{-2}$. Also plotted in this figure is the rescaled period of the oscillations. The time evolution of the injected current during two cycles of the oscillation is given in figure 3 for $\kappa/J_e^{1/2} = 4.47$ and several values of α. Here the time has been normalized with the period of the oscillation to make the different curves comparable. As can be seen, the current is nearly sinusoidal when α is only slightly above its critical value, but develops sharp rises and tall peaks separated by long intervals of very weak injection when α increases. The space charge distribution (not displayed) changes with increasing α from a slightly inhomogeneous streak to essentially independent lumps. Typically two lumps are seen simultaneously in the computational domain used.

This work was supported by DGES grant PB98-0142-C04-04.

References

1. A. Castellanos: IEEE Trans. Electr. Insul. **26**, 1201 (1991).
2. P. Atten: IEEE Trans. Diel. Electr. Insul. **3**, 1 (1996).
3. L. Onsager: J. Chem. Phys. **2**, 599 (1934).
4. M. Silver: J. Chem. Phys. **42**, 1011 (1965).
5. G. Briere, F. Garpard: J. Chim. Phys. **64**, 1071 (1967).
6. D. F. Blossey: Phys. Rev. B **9**, 5183 (1974).
7. A. Denat, B. Gosse, J. P. Gosse: J. Electrostat. **7**, 205 (1979).
8. N. Felici: J. Electrostat. **12**, 165 (1982).
9. R. Tobazeon, M. Haidara, P. Atten: J. Phys. D: Appl. Phys. **17**, 1293 (1984).
10. M. Haidara, P. Atten: IEEE Trans. Indust. Appl. **21**, 709 (1985).
11. P. Atten, M. Haidara: IEEE Trans. Electr. Insul. **20**, 187 (1985).
12. T. Takashima, R. Hanaoka, R. Ishibashi, A. Ohtsubo: IEEE Trans. Electr. Insul. **23**, 645 (1988).
13. A. Denat, J. P. Gosse, B. Gosse: IEEE Trans. Electr. Insul. **23**, 545 (1988).
14. J. L. Hernández-Ávila, N. Bonifaci, A. Denat: IEEE Trans. Diel. Electr. Insul. **1**, 412 (1994).
15. F. J. Higuera: Phys. Fluids, submitted (1999).
16. F. J. Higuera, P. A. Vázquez: Phys. Fluids **11**, 2434 (1999).

Science in the Sandbox: Fluctuations, Friction and Instabilities

Robert P. Behringer[1], Eric Clément[2], Junfei Geng[1], Dan Howell*[1], Ljubinko Kondic[3], Guy Metcalfe[4], Corey O'Hern**[1], Guillaume Reydellet[2], Sarath Tennakoon[1], Loic Vanel**[2], and Christian Veje†[5]

[1] Department of Physics and Center for Nonlinear and Complex Systems, Duke University, Durham NC, 27708-0305, USA
[2] L. M. D. H.-UMR 7603, Université de Pierre et Marie Curie, Boite 86, 4 Place Jussieu, F-75252, Paris, France
[3] New Jersey Institute of Technology, Mathematics Department, Newark, NJ 07102, USA
[4] CSIRO, Melbourne VIC 3190, Australia
[5] PMMH E.S.P.C.I. 10 rue Vauquelin, 75231 Paris Cedex 05, France

Abstract. The study of granular materials is a novel and rapidly growing field. These materials are interest for a number of reasons, both practical and theoretical. They exhibit a rich of novel dyanamical states, and they exhibit 'phases'–solid, liquid, and gas–that resemble conventional thermodynamic phases. However, the presence of strong dissipation through friction and inelasticity places these systems well outside the usual class of systems that can be explained by equilibrium thermodynamics. Thus, there are important challenges to create new kinds of statistical physics and new analytical descriptions for the mean and fluctuating behavior of these materials. We explore recent work that focuses on several important issues. These include force propagation and fluctuations in static and driven systems. It is well known that forces propagate through granular structures along networks–force chains, whose structure is a function of history. It is much less clear how to describe this process, and even what kind of structures evolve in physical experiments. After a brief overview of the field, we consider models of force propagation and recent experiments to test these models. Among the latter are experiments that probe force profiles at the base of sandpiles, and experiments that determine the Green's function response to point perturbations in granular systems. We also explore the nature of force fluctuations in slowly evolving systems, particulary sheared granular systems. These can be very strong–with rms fluctuations in the force that are as strong as the mean force. Finally, we pursue the analogy between conventional phases of matter, where we particularly focus on the transition between fluid and solid granular states in the presence of sustained horizontal shaking.

1 Introduction

Granular systems have captured much recent interest because of their rich phenomenology and important applications[1]. Despite their mundane appearance, granular materials are at the heart of many industrial processes, and are associated with a major portion of the economy[2]. These materials also appear routinely in nature. Catastrophes associated with granular materials are only too

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 351–391, 2001.
© Springer-Verlag Berlin Heidelberg 2001

common, both in industrial contexts[19], and in such phenomena as mudslides, avalanches and other destructive natural events that often lead to loss of life and property damage.

Granular materials are collections of macroscopic particles that interact via dissipative contacts. Thus, a granular system will lose all its kinetic energy in a short time, either through frictional interactions or through inelastic collisions or both. The typical energy associated with moving or accelerating a single grain (say in a gravatational field) is vastly greater than thermal energies, $k_B T$. Thus, granular materials are a-thermal: energy is lost by dissipation to heat which is more or less irrelevant to the properties of the grains, and the thermal energy bath does not couple to the motion of the grains.

Although granular materials are not thermodynamic systems, they nevertheless display different 'phases' that resemble liquids, dense fluids, and gases[1]. In this work, I will focus on the denser phases, granular solids and fluids. However, there has been much recent work on the gas-like phase as well[4,5].

The salient features of each of the granular phases are roughly as follows. In the solid state, the stress is carried on a spatially tenuous network of stress chains, as in Fig. 1. Such networks have a complex and at best partially understood structure. The contacts between grains are long-lived, and typically involve friction. In the dense fluid-like state, spatio-temporal fluctuations can be very large, with a finite probablility that there is a local fluctuation exceeding the mean by an order of magnitude or more. In the gas-like regime, models[4] based on analogies to kinetic theories are reasonably successful in predicting some, but not all properties, such as the pressure, viscosity or 'thermal' conductivity[5]. Although granular systems are a-thermal in the sense that energy transfer between true thermal modes of the solid particles is irrelevant to the state of the system, in the case of granular 'gases', it is possible to define a 'granular temperature' that is given by the rms of the fluctuating part of the particle kinetic energy: $T_g \equiv < (v - \bar{v})^2 >$.

Although there are parallels between the granular phases and their thermodynamic counterparts, there are fundamentally important differences too. The key feature that separates granular and thermodynamic systems is the presence of dissipation, through inelasticity and friction in the granular case. For the gas-like phase, inelasticity of the collisions is thought to lead to a spatial inhomogeneity, known as clustering[6]. Also, in this phase, a number of recent studies have shown some agreement, but also some deviations from the analogy to thermodynamic gases[4] (via Chapman-Enskog theories). For instance, typical velocity distributions are often roughly gaussian, but the exponent of the velocity may not be exactly 2, equipartition is often violated, and some transport properties are far from the predicted value[5].

On laboratory scales of length and time, fluctuations in the denser phases can be much stronger than in conventional thermodynamic systems, and we are only beginning to get a handle on their properties[7,8,9,10,11]. The role of friction in this phase is still an open question[13]. Typically, the mean forces and the rms of the force fluctuations are comparable in magnitude. The strength of these fluctu-

ations makes it difficult to obtain the appropriate long-wavelength dynamics for dense granular systems. At present, it is not clear under what conditions or scales fluctuations cease to be important, and when/if a long-wavelength description is guaranteed to work. Perhaps the most fundamental unresolved issue for dense granular systems is the determination of the appropriate statistical description. Necessarily, this description is different from the statistical mechanics of thermal systems, because ergodicity often does not apply. Thus, it is possible that we may be able to predict the probability of a given force distribution within a granular system. However, it is an experimental fact that various realizations of a flow or even a static configuration of grains developed under nominally identical macroscopic conditions can show significant differences.

Even in the absence of strong spatial disorder of the grains, dense granular arrays show inhomogeneous spatial stress profiles called stress chains[14], where forces are carried primarily by a small fraction, perhaps a third, of the total number of grains (see Fig. 1). Under slow deformation, such as that which occurs due to shearing, changes in the stress chains lead to the strong temporal fluctuations[8,10] noted above. We have explored this feature in several different experiments which we discuss below.

There are two concepts that are particularly relevant to deformation of dense granular materials: Reynolds dilatancy, and Coulomb friction. If shear is applied to a densely packed sample of hard frictional grains, the system must dilate (Reynolds dilatancy) so that the grains can move around each other[15]. Under sustained shearing, the dilation and displacement of grains is often limited to a region that is $\sim 5 - 10$ grains wide called a shear band. The second concept, Coulomb friction, is actually two concepts: first, solid-on-solid (S-o-S) friction, which was first considered by Amontons[16]; and second, Mohr-Coulomb failure, which refers to the onset of slipping within a granular solid. The union of Reynolds dilatancy and granular friction rules is at the heart of modern soil mechanics[17,18].

S-o-S friction, which is still the topic of modern research[19], is epitomized in the introductory physics problem of a block on an inclined plane: the block will slide when the tangential forces F_T exceed μF_N, where F_N is the normal force, and μ is the coefficient of static friction. In the granular analogue, slipping or Mohr-Coulomb failure, occurs along surfaces where the ratio of shear and normal stresses exceeds the corresponding friction coefficient. The interesting (and a bit murky) aspect of friction in granular materials is that individual grains interact through S-o-S friction, but in addition, onset of failure within a granular solid is also reasonably well described by Mohr-Coulomb frictional rules.

It is important to note that S-o-S friction within a collection of granular particles is indeterminate for all but very simple cases. This is demonstrated in Fig. 2, which contrasts three cases: a) a block on a plane, b) two disks on a plane, and c) three disks on a plane. Assuming a static state, it possible to determine all the frictional forces in cases a and b, but not for c, or more generally for multi-paricle states. Specifically, for case c, there is one more unknown force component than there are constraints, so there is a one-parameter family of

solutions that satisfy force and torque balance. The addition of more particles typically leads to additional undetermined degrees of freedom, depending on the number of contacts added. In the absence of friction, i.e. if only normal forces are allowed, the forces are uniquely determined. When all there are exactly the same number of constraints and degrees of freedom, this is refered to as isostaticity[20]. To the extent that real granular systems involve fricitonal forces, the resulting indeterminacy of static granular arrays, implies that these systems are sensitive not just to the external applied constrains, but also to the way in which the system evolved. The challenge is then to relate control parameters and history to the statistical and mean properties of the system.

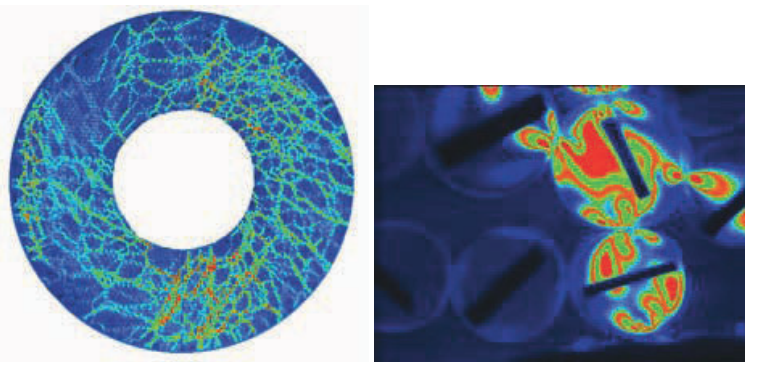

Fig. 1. Left: Image of stress chains from a shear experiment discussed below. The stress chains are visualized by photoelasticity: brighter regions correspond to larger forces. Right: Close-up image of several disks showing fringes or bands. Note also dark bars used to track individual disks.

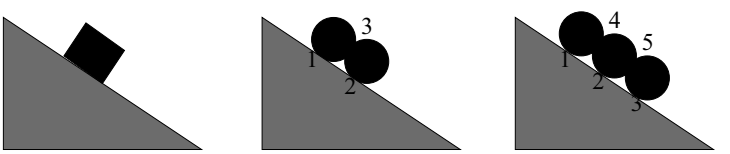

Fig. 2. Sketch of three situations: a) a block on a plane, b) two disks in contact on a plane, and c) three disks in contact on a plane. In a) all forces are known. In b) there are 6 unknown force components, and 6 constraints (2 force components sum to zero and 1 torque component sums to zero per disk), so all forces are again known. But, in c) there are ten unknown force components and only 9 constraints.

Recent work[21] also suggests that granular systems may be part of a broader class of systems, known as jammed systems. Among the members of this class,

in addition to granular materials, are fluids near the glass transition, foams, colloids, and suspensions. The unifying feature of these systems is that they undergo a transition from a mobile to a constrained state or vice versa as some parameter or set of parameters changes. In the case of the glass transition, the relevant parameter is the temperature, T, and for granular materials the relevant parameters include both stress and density. That is, under shear, dense granular systems dilate (change density) and become mobile.

Several recent theoretical works[7,22,23,24,25] have provided a context for understanding static stress distributions in granular materials. The first of these, the q-model[7] of Coppersmith et al, predicts a force distribution for static systems $P(F) \propto F^{N-1} \exp(-F/F_o)$, where N is the system dimension. This model assumes a regular array of grains, nominally subject to gravity, neglects torques, and considers the force balance in one direction only. It introduces randomness by assuming that the force on a grain in one layer is balanced by transmiting fractions q_i, with $\Sigma_1^N q_i = 1$, to the N supporting grains in the next layer of an ordered dense packing, assuming an N-dimensional system. The q_i are random numbers uniformly distributed in $0 \leq q_i \leq 1$, with the summation constraint above. For concreteness, it is useful to imagine the 2D case, where the 'grains' are disks. Then, a fraction, q and $1 - q$ of the force on a grain is transmitted to the two lower neighbors. In addition to an exponential distribution of large forces, the q-model predicts that forces between neighboring grains are uncorrelated and that in the long-wavelength limit, the stresses follow a diffusion equation, with the time axis along the direction of gravity. Various other lattice models[22,23,24] contain features that guarantee torque and vector force balance, the possibility of correlated stress chains, or other features of random packing.

A rather different approach is embodied by the Oriented Stress Linearity (OSL) model of Bouchaud, Cates, Witmer and Claudin[26]. The OSL picture supposes an anisotropic friction-like consitutive law for the components of the stress: $\sigma_{xx} = \eta\sigma_{zz} + \mu\sigma_{xz}$, that is based on an intuitive picture of the robustness of stress chains under different directions of applied stress. The quantities η and μ are constants that are expected to depend on the history of the sample preparation. When this consitutive rule is combined with the conditions of stress balance: $\nabla_i \sigma_{ij} = \rho g_j$, where ρ is the material density and g_i is the appropriate component of the gravitational acceleration, the result is a purely wave-like hyperbolic equation for stress propagation throughout the material: $(\nabla_z + c_+\nabla_x)(\nabla_z + c_-\nabla_x)\sigma_{ij} = 0$. The creators of the OSL model have also discussed how such a continuum limit might occur in the context of a lattice model that they dub the Three-Leg Model[26].

The contrast between the q-model and the OSL model is dramatic. In the long-wavelength limit of the q-model, information is lost rapidly through diffusion, whereas in the OSL model, information propagates through the medium. The difference in the models is also apparent in the effects of the boundaries on the possible solutions. For instance, less boundary information is needed for a hyperbolic equation than for an elliptic equation.

Kenkre et al.[27] have proposed a continuum model that can, in essence, interpolate between the q-model and OSL. The overarching principle of their model is the idea of a memory function, such that the response of the granular material is obtain as a convolution of applied stresses with the memory function. Either a diffusion equation, a wave equation, or the telegraph equation results from an appropriate from the memory function.

An alternative approach, contact dynamics[25], has been developed by Moreau, Radjai, and Roux. This scheme provides an algorithm to lift the frictional degeneracy of the contacts for hard frictional particles. These authors have considered a number of different granular systems, included systems under compression and shear. For the present purposes, we note that this model also predicts $P(F) \propto \exp(-F/F_o)$ for large forces.

2 Experimental Tests of Static Models

There have been several attempts recently, some ongoing, to test predictions of the models described above. We will focus on three of these, 1) tests of the static force distributions, 2) measurements of force profiles under sandpiles, and 3) measurements of the Greens function, i.e. the stress response function of a granular material to a local force.

2.1 Force Distributions

An important measure of the statistical properties of granular materials is $P(F)$, the distribution of forces on individual particles. Both the q-model, a number of related lattice models, and contact dynamics, make the striking prediction that this distribution will fall off exponentially at large F, although the predictions do not agree about what happens for small F. For instance, in the q-model at small F, $P \propto F^{N-1}$ where N is the dimension, and in contact dynamics simulations, the exponent $N-1$ is replaced by another exponent that is nearly zero.

The large-F prediction has been tested by the University of Chicago group, using a novel 'carbon paper' technique[7,9]. The forces at the boundaries of a container of material were measured by placing a sheet of carbon paper against the boundary. The size of the mark left after compaction was a monotonic function of the applied force. From these measurements, the authors showed clearly that the contact forces followed an exponential distribution at large F.

Several other groups have also reported an exponential distribution, but typically in time-evolving rather than simply static systems[8,10,11]. In both 2D and 3D Couette-like sheared systems[8,11] (discussed below), we have observed this large-F exponential distribution. (See also Shiffer et al. who have carried out experiments in which a rod was driven through a granular sample[12].)

Thus, an exponential form for the large-F part of the distribution seems to be a recurrent theme. A less well established issue is the nature of the small-F part of the distribution, and the nature of spatial correlations. We discuss this further below.

2.2 Force Profiles under Sandpiles

The OSL model has been used to interpret a classic problem in granular mechanics, specifically, the force profile at the base of a granular heap, or sandpile[29,26]. Experimenters have variably reported that the force under a conical heap of material has a minimum or a minimum under the peak of the pile (see Fig. 3). A minimum under the tallest part of the heap is counterintuitive, since we normally expect that forces propagate hydrostatically. A key insight of the OSL model is to note that stresses can propagate through stress chains much as stresses propagate through butresses of a cathedral. If the history of the heap preparation leads to organized arches down the slope of of the heap, then the force might be a maximum on a ring of radius $R_m > 0$. The OSL model was constructed in such a way that this property was manifest, depending on the choice of the parameters η and μ.

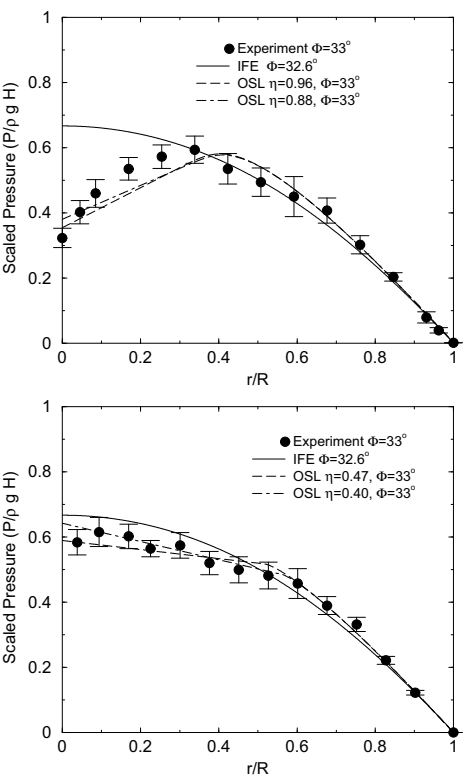

Fig. 3. Data for the force at the base of a sandpile vs. distance from the center for top: piles formed by pouring from a localized source–a funnel, and bottom: piles formed by pouring from a distributed source–a seive. After Vanel et al.

In yet a different direction, Savage[28] has pointed out that classical soil mechanics models, based on the concept of incipient failure everywhere (IFE), can also explain differences in the force profiles under heaps. The idea is that the inner layers of the pile undergo an elastic deformation because of gravity, but that the outer part of the heap is arbitrarily close to slipping, i.e. Coulomb failure. Savage also pointed out that past experiments may be subject to significant errors because even very tiny deflections of the surface on which the piles form are sufficient to disturb the stress chains/stress profiles.

Several experimental groups have set out to explore this issue recently[29,30]. Here, we focus on the work of Vanel et al[29]. Using very sensitive transducers and very stiff surfaces, we found a clear variation of the force profile under sand heaps, depending on how the heap was prepared. If the heap was poured from a funnel, thus creating arches down the slope of the heap, there was an undebatable minimum at the center of the heap. If the heap was prepared by pouring from a sieve, there was a weak maximum at the the center of the heap. We contrast the force profiles for these two methods in Fig. 3. Similar, although not as dramatic, results occur for wedge-shaped piles. Recently, we have pursued the issue of history dependence in the context of 2D 'sandpiles' using photoelastic disks[11] (photoelasticity allows us to measure local forces on the disks, as discussed below). Using preparation techniques that are the 2D analogies of funnels and sieves, we find that the most dramatic difference occurs in the structural nature of the piles, as evidenced in Fig. 4. This figure shows the contact angles between nearest neighbor disks in mutual contact. In the case of the pile formed from a funnel, there is strong anisotropy; in the case of the pile formed from a sieve, there is much weaker anisotropy of the grain contact angles.

2.3 Greens Function Measurements

There is a diverse spread in predictions for the large-wavelength behavior of static granular systems. Thus, the q-model predicts diffusion–hence a parabolic PDE; the OSL model predicts a wave equation, hence a hyperbolic PDE, and the occurence of 'light cones'. Conventional models predicts an ellasto-plastic response, hence for part of the material an elliptic PDE. The models of Kenkre et al. provide both a hyperbolic interpolation between the parabolic and wave cases, and more generally, a mechanism to examine other possible response functions.

This means that a fundamentally important experiment to carry out is the determination of the Green's function, i.e. the response of a granular system to a localized point-like perturbation. Ideally, we would like to know in detail the change in the forces on each grain as a function of position, following the addition of the perturbation. The average over an ensemble of realizations is then a representation of the continuum behavior.

We are carrying out such experiments on a 2D sample consisting of photoelastic particles. The ideas behind the technique are simple: photoelastic materials are birefringent under stess. Because the particles are photoelastic, we can obtain information on the stress state of the particles by means of a simple polarimeter. As polarized light traverses a photoelastic material, there is a

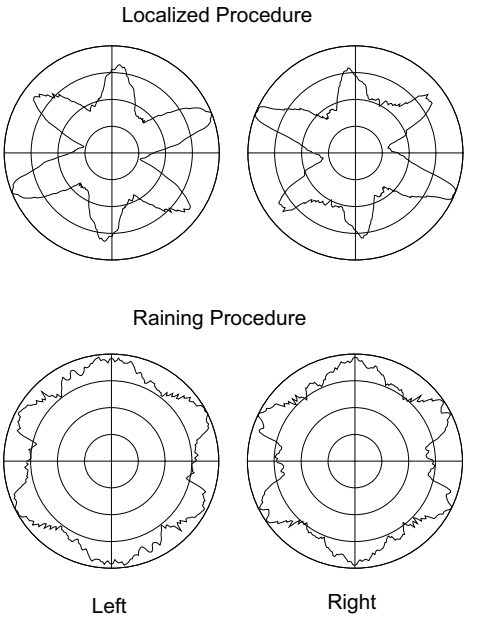

Fig. 4. Distributions of the contact angles for 2D piles formed by pouring from a funnel (top) and a seive (bottom). For each preparation method, we show the contact angle distributions for each side of the final heap. It was essential to distinguish right and left, since the combined distribution is not sensitive to the 6-fold orientation of the contacts.

polarization-sensitive phase shift (birefringence) depending on the difference in principal stresses, $\sigma_2 - \sigma_1$ in the material along the path of the light. Particles in a polarimeter[31], (see inset of Fig 19) produce images with light and dark bands or fringes, as in Fig. 1 (Right), which shows a closeup of a photoelastic image of several disks. These fringes, corresponding to successive wrappings of the polarization angle of the transmitted light, increase in number with the local force at a grain contact. Although it is possible to determine the stress state of the disks[32] by examining the geometry of fringes, the level of complexity, resolution, and computational time, is demanding if large amounts of data are to be acquired. We take a different and novel approach that exploits the fact that the number of fringes increases with the local contact force. This means that the gradient of the image, averaged over a disk size, grows with the applied forces on all the contacts. Thus, we calculate $G^2 \equiv < |\nabla I|^2 >$, where I is the pixel intensity and the average is computed over the scale of a disk. We make this procedure quantitatively accurate via a calibration such as Fig. 23, Inset. To obtain this calibration, we apply known forces to the perimeter of a modest number of disks. Note that G^2 is a monotonic reproducible function of the force per perimeter length, independent of the number of disks in the calibration sample.

The results presented here are preliminary, but, they are nevertheless very interesting, and point to a number of different questions. A layer of photoelastic particles (more information given below), either pentagons or disks, is oriented nearly vertically. We record the photoelastic image before and after the application of a point force, and we then compute the difference in stress due to that force. It is the difference in these images, Fig. 5, that is important.

It is very clear from this last image, that the details of the response depend very much on the specifics of how the grains are arranged. If we repeat the observation with a different configuration of grains, we obtain a different response. Our expectation is that the ensemble average of nominally identical systems will yield the appropriate mean response to which we might compare a model. Since each realization for a given configuration is specific and complex, this seems like the only way in which a modest scale experiment can probe this issue.

We show the average response (the difference in $G^2 \equiv |\nabla I|^2$ averaged over multiple trials in Fig. 6, with the left image corresponding to the response of a packing of identical disks (minimal packing disorder), and the right corresponding to the response of pentagonal particles (no spatial ordering expected). Although there are some difference in the two types of response, there also several qualitatively similar features: 1) the presence of an outer envelope of roughly linear form, 2) peaks that evolve and broaden as the distance from the source increase. Fig. 7 shows horizontal cuts across these two images that show this information more quantitatively. These experiments suggest a more complex response, such as dispersive wave propagation, rather than that contained in any of the models. However, we emphasize that these are still preliminary results.

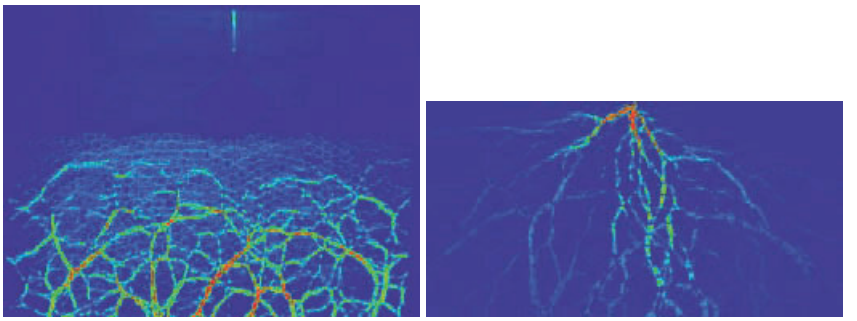

Fig. 5. Left: Photoelastic response before the application of a point force. Right Single realization of the photoelastic response of an applied point force. Both for pentagons.

3 Fluctuations

Until recently, the issue of fluctuations in dense granular systems was relatively unexplored. However, two experiments[33,34] suggested that such effects might

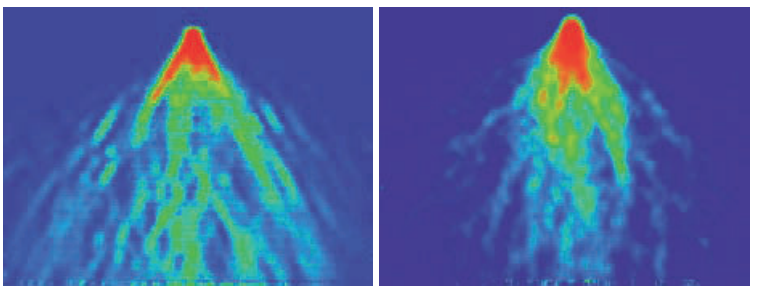

Fig. 6. Averages over many (typically 50) different realizations of the photoelastic response of an applied point force for left: disks; and right: pentagons.

be very important. One of these, stress measurements for a granular material that was flowing from a hopper[33] showed unexpectedly large scale non-periodic fluctuations with spectra that were roughly power-law like. Here, the measurements were performed at the inner wall of the hopper, and away from the rapidly moving material near the exit of the hopper. Another experiment of interest involved sound propagation experiments by Liu and Nagel[34]. In these experiments, a local oscillating force was applied to a spherical particle, and the response was detected some distance away with a microphone. A particularly striking feature of these experiments was the very complex structure seen in the response amplitude as a function of frequency (see also Jia et al.[35]). This structure was highly sensitive to even very tiny displacements of a grain.

More recently, we have carried out experiments explicitly designed to characterize fluctuations in both 2D and 3D granular systems[8,36,37,11,38]. In the discussion below, we review some of the features of each. Perhaps one of the most striking results of the 2D experiments is the characterization of a continuous transition as the density, or packing fraction is varied through a critical value, γ_c, such that stresses can just be sustained in the material.

3.1 Slow Shearing in 3D

In 3D[8], we slowly sheared a sample of spherical glass particles in an annular container that is sketched in Fig. 8. Shearing was achieved from above by a heavy rotating ring that was roughened by a layer of similar spheres, glued to the ring. This ring was free to translate vertically and was driven azimuthally at constant angular velocity, ω_r, where ω_r was varied over $0.003 \leq \omega_r \leq 1\,\mathrm{rad/s}$ in different experimental realizations. A fast sensitive capacitative force sensor that was flush with the bottom of the shear cell, yielded σ, the normal stress, vs. time on a cross sectional area of 0.8 cm. Perhaps the most striking aspect of these experiments is that the stress fluctuations could be much larger than the mean stress in the system, as seen in Fig. 9.

In order to better understand correlations and stress propagation effects, we varied the height of the layer, h, and the diameter of the spheres d, with

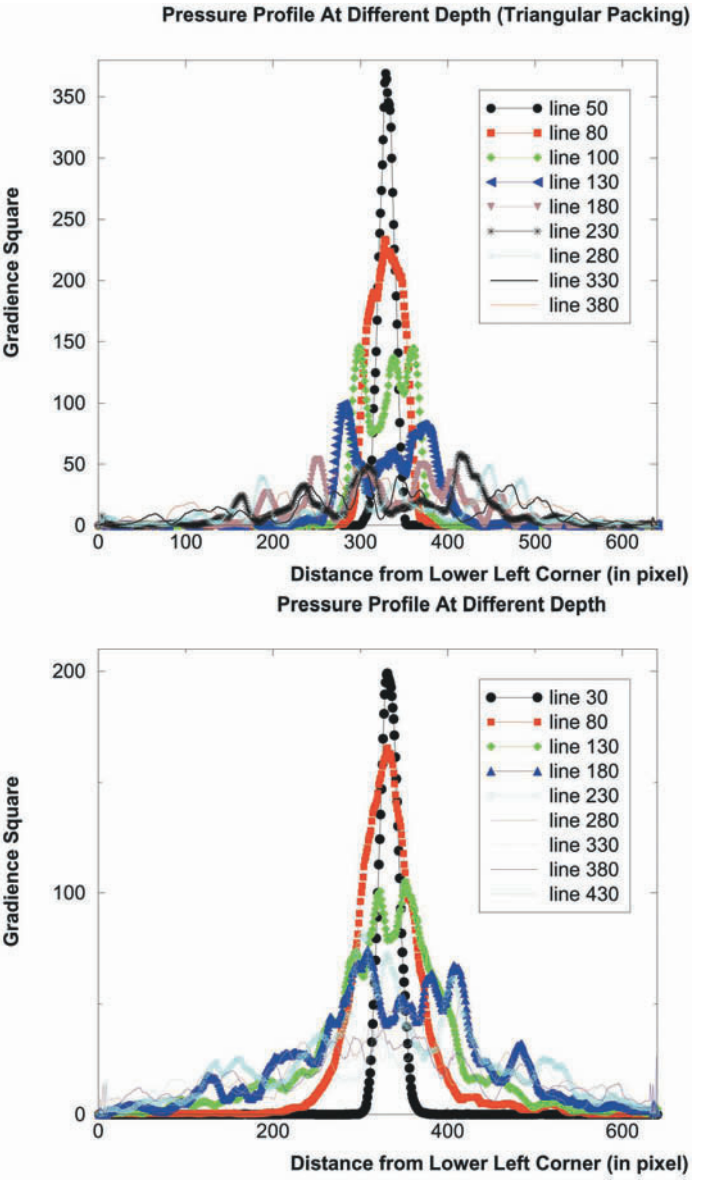

Fig. 7. Quantitative realizations of the previous figure: above: disks; and below: pentagons. Each curve is the value of the difference in G^2 along horizontal lines transverse to the direction of the applied load.

$1.0\,\text{cm} \leq h \leq 4.1\,\text{cm}$, and $0.1\,\text{cm} \leq d \leq 0.5\,\text{cm}$. Since the area of the detector was fixed, by changing the sphere diameter, we were able to change n, the number of spheres contacting the detector from ~ 4 to ~ 100. If the forces on nearby grains were uncorrelated, then we would expect to see a narrowing of the distribution

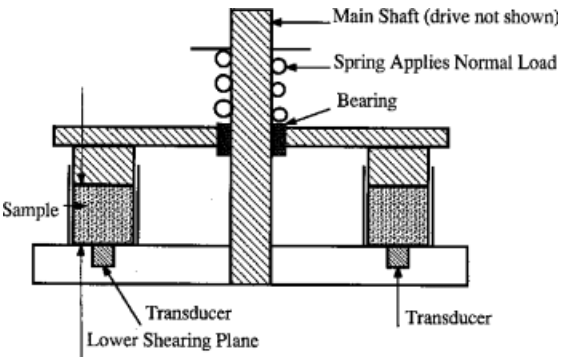

Fig. 8. Schematic of apparatus used for 3D shearing experiments. The figure is a vertical cross section of a system with cylindrical symmetry (excepting the transducers). Glass spheres lying in an annular channel are sheared from above by a rotating ring on which spheres are glued. Normal stresses are measured vs. time at the base of the channel by the transducers.

by $n^{-1/2}$, the number of particles contacting the transducer, i.e. the distribution would be narrowest for the smallest particle diameters.

Returning to the time series of Fig 9, we note that, the probability of having a fluctuation an order of magnitude greater than the mean is non-negligible (see below). An important question that remains to be addressed is what happens as the system size becomes very large.

Power spectra resulting from such stress time series provide useful information. Fig. 10, left shows spectra at various rotation rates, ω_r, of the shearing wheel, where the particle diameter is $d = 2.0\,\mathrm{mm}$. At high spectral frequency, ω, P falls off as $P(\omega) \propto \omega^{-2}$ for a given ω_r. The spectra vary more weakly with ω at low ω, but in this range, $P(\omega) \propto \omega^{-\alpha}$, with $0 < \alpha < 2$. The high frequency dependence is easily explained by assuming a series of random jumps in $\sigma(t)$ occuring at least as fast as a crossover time $\sim 1/\omega^*$, corresponding to a shearing ring displacement of ~ 1 particle diameter. The low frequency behavior is perhaps more interesting, since $\alpha > 0$, would imply processes at slow time scales. A possible source for such a slow process may be long-time rearrangements of the grains.

These spectra also show a kind of rate invariance. To place this in context, many continuum models of slow granular flow[39,18] assume rate invariance for the mean flow properties. The rate invariance shown here is interesting because it applies to the *fluctuations*. To interpret this result, we note that if the grains were almost always close to static equilibrium, we would expect that the time series for the measured force fluctuation would have the form $F(t) = g(\omega_r t)$. Here, $g(z)$ might depend on the history of the system, but not on ω_r except through the term $\omega_r t$. If we sampled over a long time interval, T, any history dependence would be removed as the system explored a large collection of the possible nearly

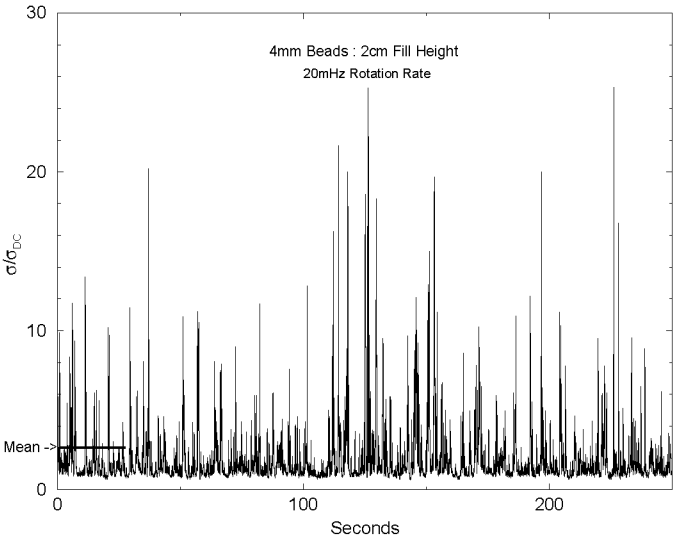

Fig. 9. Time series showing stress fluctuations for grains sheared in the 3D annular channel. Stresses are normalized by the static stress, σ_{DC}, measured at the beginning of the run. The small bar at the lower left indicates the mean stress.

quasi-static states. In that event, a power spectrum for the fluctuations would have the form

$$P(\omega) = T^{-1}|\int_0^T \exp(i\omega t)F(t)dt|^2 = (\omega_r \bar{T})^{-1}|\int_0^{\bar{T}} \exp(iz\omega/\omega_r)g(z)dz|^2, \quad (1)$$

where $\bar{T} = T\omega_r$, and $z = \omega_r t$. Eq. 1 indicates that the scaled power, $P\omega_r$, whould depend only on the dimensionless spectral frequency ω/ω_r. Alternatively, ω_r sets the time scale for the fluctuations. This expectation is borne out in these experiments, as typified by Fig. 10, right. This figure shows an excellent collapse of all the data for the scaled power vs. the scaled frequency for two different particle sizes. Similar collapse was found in all cases. This result is interesting for at least two reasons: first because this is the first demonstration to our knowledge of rate invariance for the fluctuations, and second, because data for two very different size grains show nearly the same scaled power spectra.

A final interesting aspect of these experiments is the distributions of forces (stresses) without regard to time. Thus, the shearing creates a sequence of stress that are members of an ensemble. Since that the time series show rate invariance, it is meaningful to compare the measured force distributions to models and experiments for static situations. Fig. 11 shows such distributions for four different particle sizes. For the smallest (1 mm) particles, roughly $n = 100$ particles contacted the stress transducer; for the largest particles, roughly $n = 6$ contacted the transducer. For all cases, the distributions are roughly exponential

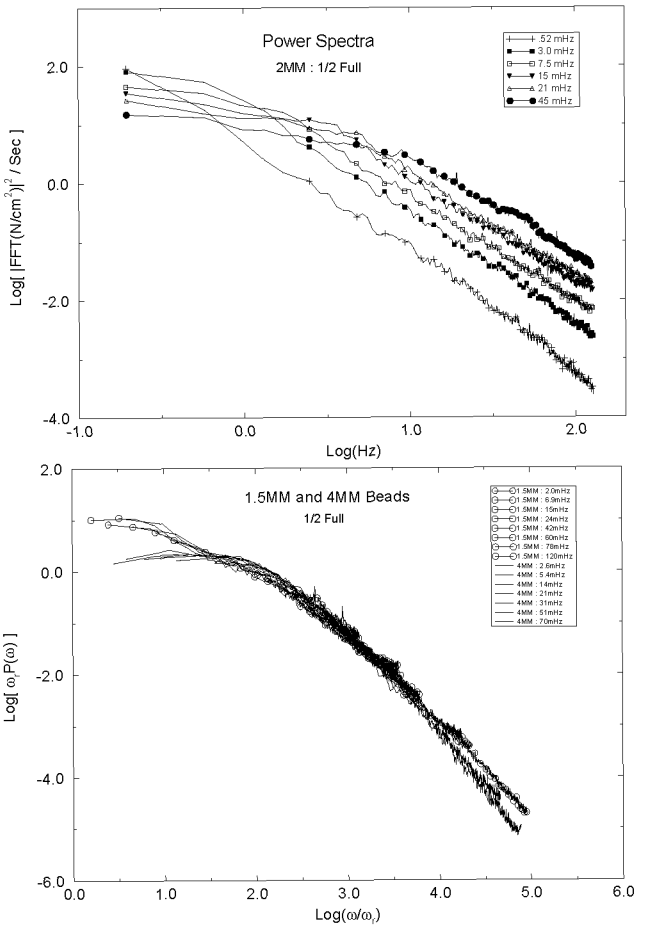

Fig. 10. Above: Typical power spectra for normal stress fluctuations in the 3D shearing experiments for 2 mm particles that half-fill the 4 cm deep channel. Rotation rates, $\omega_r/2\pi$ are indicated in the box. Below: Scaled power spectra, $\omega_r P(\omega/\omega_r)$ vs. ω/ω_r for normal stress fluctuations in the 3D shearing experiments. Results are given for two different sphere sizes and for the indicated rotation rates of the shearing ring. The 4 cm deep channel was filled to a depth of 2 cm in these experiments.

for large forces, as in the models discused above. Interestingly, the distributions of Fig. 11 do not show the $n^{-1/2}$ narrowing that would occur if the force on each grain were totally uncorrelated with its neighbors. Thus, for n uncorrelated grains described by the q-model, the total force F, summed over these grains is distributed as $\rho_N = [f_o(3n-1)!]^{-1}(F/f_o)^{3n-1}\exp(-F/f_o)$, where f_o is an appropriate normalization constant. The multi-particle distribution of forces for n particles at a time then narrows as $n^{-1/2}$, as sketched in Fig. 12. For the experimental distributions in Fig. 11, the mean number of particles in contact with

the transducer changes by a factor of 16 as the particle size changes from 4 mm to 1 mm. Although the distributions change qualitatively over this size range, they do not show a narrowing by a factor of 4, as would occur in the absence of correlations. Within the scatter, there does not appear to be as strong a narrowing as one would expect for totally uncorrelated forces. Interestingly, in the carbon paper experiments of Mueth et al.[9], no spatial correlations were found. This difference in the two experiments is unresolved at the moment.

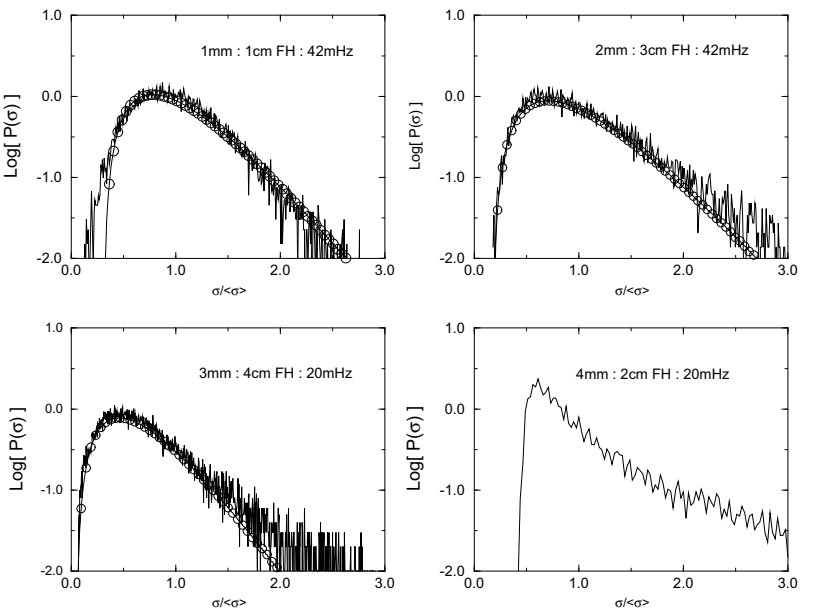

Fig. 11. Stress distributions, $\rho(\sigma)$, from the 3D shearing experiments for four different particle sizes. A small positive quantity has been added to ρ to ensure that the logarithm always exists.

3.2 Slow Shearing in 2D

Experiments in 3D cannot provide detailed information about the internal forces or movement of the particles. However, this kind of information *is* available via 2D measurements in which the particles are constructed of photoelastic material. Thus, in a second set of experiments[37,40,11], we have characterized fluctuations in a 2D granular system consisting of either circular disks or pentagonal particles confined to a plane and subject to shearing. In this system, we were able to measure forces, displacments, and rotations of the individual particles.

The experimental arrangement is sketched in Fig. 13. The disks slide on a smooth powder-lubricated sheet of Plexiglas. They are confined radially by a

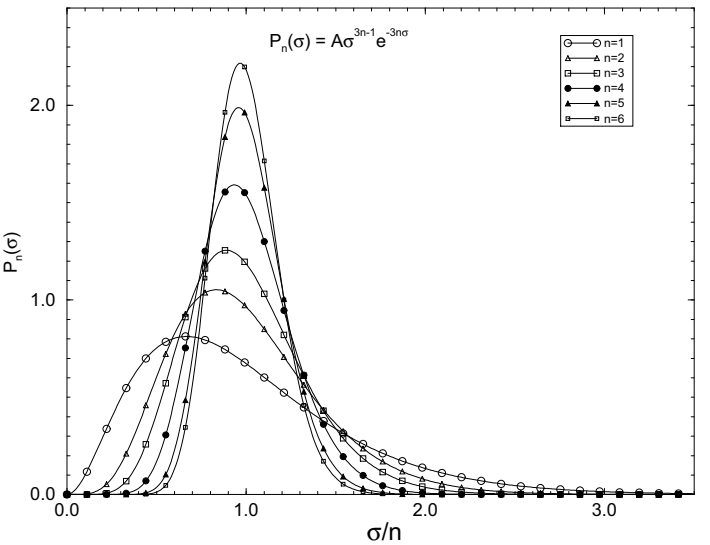

Fig. 12. Narrowing of force distributions for collections of N uncorrelated grains. The distributions are calculated for the collective force of N grains satisfying the 3D q-model single-particle distribution $\rho = (2f_o)^{-1}(f/f_o)^2 \exp(-f/f_o)$.

wheel at the inner radius, $r = R_i$, and by a ring at the outer radius $r = R_o$, where the surfaces of the wheel and ring contacting the grains are roughened, as sketched in Fig. 13. The disks used here are the same as the ones used for the Greens function experiments. They are cut from a photoelastic[31] polymer, which allows us to determine information on the stress. The diameters, d, of the disks are either 0.9 cm or 0.74 cm, with $\sim$ 400 of the former and $\sim$ 2500 of the latter. All disks are 0.62 cm thick. Two sizes of disks were used to reduce the possibility of crystalline ordering. Another concern, segregation by size does not occur in these experiments. The pentagons had the same thickness, and were 0.64 cm on a side. Additional details can be found in several references[37,40,11].

Here, we consider shearing only from the inner wheel, although it is possible to shear from the outer ring as well. The rate of shearing is set by $\Omega = 2\pi/\tau$, the angular velocity of the inner wheel, where τ is the time for the wheel to complete a revolution. An additional control parameter is the density of particles, as specified by the packing fraction, γ. ($\gamma \equiv$ the area of the annulus occupied by particles normalized by the total area).

In addition to force information, we also determine kinematic properties of the system by tracking small dark bars placed on each disk (see Fig. 13). A sequence of video images is digitized by a frame grabber and analyzed to determine

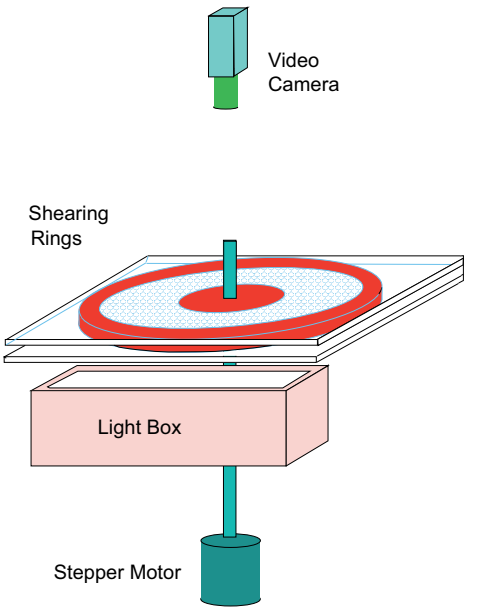

Fig. 13. View of apparatus used in the 2D shear experiments. Photoelastic disks lie between two smooth dry-lubricated Plexiglas sheets. The disks are sheared by a rotating wheel, and confined by a ring. This ring is held in place by planetary gears.

the center and orientation of each bar in each video frame. Using computer software, we track individual particles from frame to frame in the video to obtain the velocity, $\boldsymbol{V}$, and angular velocity or spin, S, of each particle. From samples over long time periods, we construct the statistical and mean kinematic properties of the system as a function of the radial distance r.

We separately determine the local packing fraction, γ_l by using the fact that the disks strongly attenuate blue/UV light. We observe the transmission through the disks of light from a fluorescent "black light". With this procedure, we determine γ_l with an absolute accuracy of about 1%. However, relative changes in the global packing fraction γ are determined much more precisely, since this involves the removal or addition of a known number of particles.

Before turning to experimental results, we note that MD type simulations by the Stuttgart computation group[41,42] are in progress that replicate as closely as possible the details of these experiments. These calculations capture almost all the observed features of the experiment, and provide additional insights regarding quantities that are difficult to obtain experimentally.

An important aspect of shearing is the formation of a strongly dilated region or shear band[43], as seen in the present experiments in Fig. 14. This figure gives γ_l vs. r, the distance from the shearing wheel in units of a particle diameter. The dilated region extends out about 6 particle diameters from the wheel, and

the profile is consistent with an exponential:

$$\gamma_l(r) = A - B\exp(-r/\lambda), \tag{2}$$

where λ is a characteristic length that varies with the global packing fraction.

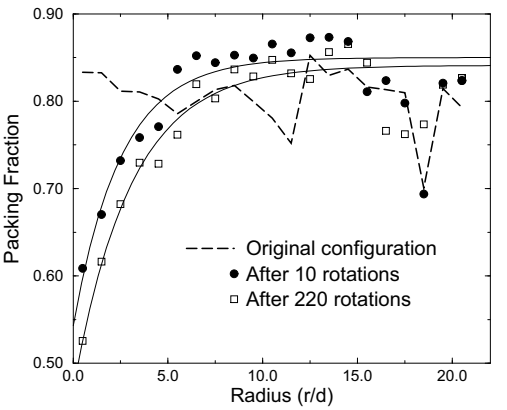

Fig. 14. Mean local packing fraction, γ_l, vs. radial distance $(r - R_i)/d$ from the wheel in units of disk diameters. The figure shows the initial profile before shearing (dashed line) the profile after 10 wheel rotations, and after 220 rotations. The mean packing fraction is $\gamma = 0.788$.

The dilated region takes a considerable time to fully evolve. This occurs for at least two reasons. First, the dilated region near the shearing wheel can only form by pushing particles outward into a more compact region, a process that becomes more difficult as that outer region becomes more compacted. This process resembles the slow dynamics that is typical of the parking problem[44] because as the system compacts, the radially displacement of a grain requires the collective displacement of numerous other grains at larger radius. Because the evolution of this band can be very slow, we typically sheared the system for $\sim 1hr$ before collecting data. A second reason for slow evolution involves the global packing fraction γ. If γ is close to a critical value discussed below, there is critical slowing down associated with the transition.

A Continuous transition in γ The statistical and mean properties of the sheared system depend signifciantly on γ. This is easy to appreciate by imagining what would happen if γ were very low: particless would initially be pushed away from the shearing wheel until no more displacements occurred. If γ were gradually increased, we might expect to find a critial value, γ_c, where there would always be at least some particles subject to compressive and shear forces from the boundaries. An important issue is the nature of this transition. We find that under the conditions of this Couette experiment, a granular system exhibits

a number of features that are similar to a continuous phase transition. At γ_c, the system is very compressible, and in a given part of the system force chains occur only intermittently. Most of the grains are close to uncompressed and static most of the time. With the addition of more grains, the system strengthens, more chains occur, and grains are dragged more frequently by the shearing wheel. A consequence of this last feature is that the system 'speeds up'. If even more grains are added, the system becomes very stiff, and disks are subject to larger deformations. In the extreme limit, brittle materials might fracture. It is this range of γ's, between γ_c and the strongly compacted regime that we explore here. It is important to note, however, that the actual change in γ is small; all the action takes place over a change in γ that corresponds to only about 5% variation.

The temporal intermittentency in the stress fluctuations near γ_c is clear from Fig. 15 which shows force fluctuations determined from G^2, where here, G^2 is integrated over about 1/10 of the system in a radial slice. For γ just slightly above γ_c (bottom time series) the system is quiescent much of the time, with sporadic bursts of activity in which stress builds up and is then released. By contrast, when the system is more densely packed (top time series) the mean stress is substantially increased, and the fluctuations occur continuously.

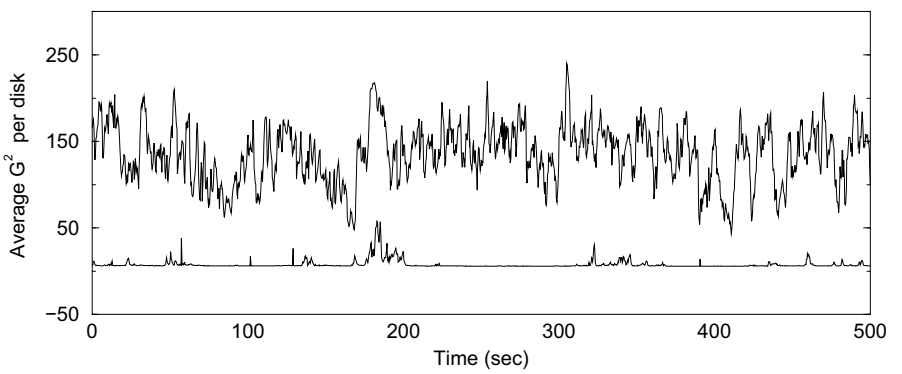

Fig. 15. Time series for stress fluctuations from the 2D Couette experiment with disks. The lower trace is just above the critical value of γ, $\gamma = 0.777$, and the upper trace is in a relatively compact state, $\gamma = 0.803$.

We characterize the intermittency of fluctuations by defining 'avalanche' events to be a negative change in G^2 exceeding the instrumental noise. (Similar data result if we look at positive changes in G^2.) Fig. 16 shows representative data for the probability distribution of avalanches on log-log and semi-log scales, for pentagonal particles (Right) and disks (left). Close to γ_c ($\gamma_c \simeq 0.73$ for pentagons; $\gamma_c \simeq 0.776$ for disks), these distributions resemble powerlaws with an exponent of about 3/2. For the denser distributions, the distributions are well described by exponentials.

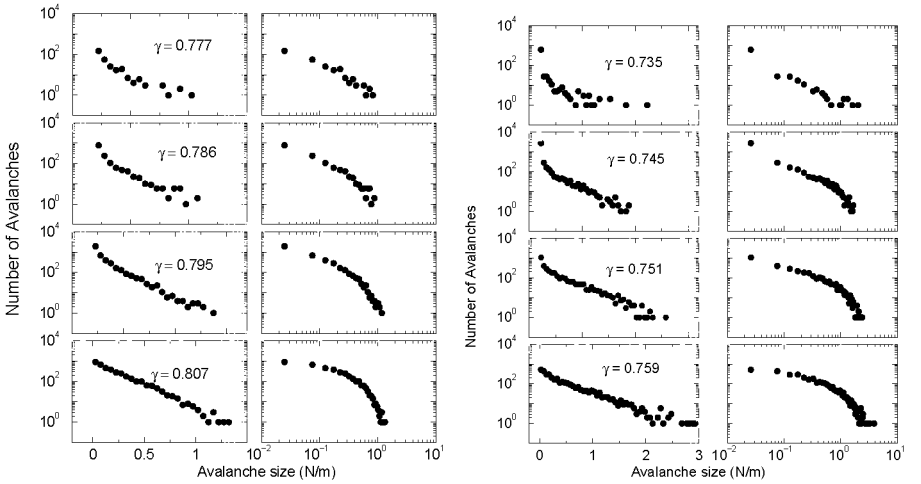

Fig. 16. Distributions of "avalanches" for representative values of γ on log-log and log-lin scales. Lefthand data pertain to disks, and righthand data to pentagons. These data provide some insight into the intermittent character of the fluctuations for $\gamma \simeq \gamma_c$. We define an avalanche in the main text.

Associated with the intermittency near γ_c is critical slowing down, as manifested in the kinematic properties, which include velocities and particle rotations (i.e. spin, S). For the present purposes, the azimuthal velocity is the most interesting because the disks are severely limited in their radial travel. We express the azimuthal motion in terms of the velocity, $V_\theta \equiv \boldsymbol{V} \cdot \hat{\theta}/r$.

The azimuthal motion occurs almost exlusively in the shear band next to the shearing wheel, as seen in data for the mean angular velocity, Fig. 17, top. We have scaled data for V_θ by Ω, the angular rotation rate of the inner wheel. If there were perfect rate invariance, for which the properties of the system depended only on the single time scale, Ω^{-1}, all the data for V_θ/Ω would collapse onto a single curve. This expectation is met reasonably well, but not perfectly, although departures from perfect rate scaling are apparently random rather than systematic. These variations may be caused by slow changes in the outer part of the system, well away from the shear band. Indeed, we observe that over very long time scales, of order hundreds of hours, such changes do occur. The data are consistent with a profile $V_\theta(r)$ of the form

$$V_\theta \propto \exp[-A - B(r - R_i)/d - C((r - R_i)/d)^2]. \tag{3}$$

These results differ somewhat from the 3D experiments of Jaeger et al. [45], who report a dominant quadratic term relative to the linear term in the exponential. The calculations of Schöllmann et al.[41,42] for this system show a profile that has an essentially linear term in the exponential.

We now return to the question of why the rate-independence is not perfect. The data suggest that motion is possible over very long time scales even at

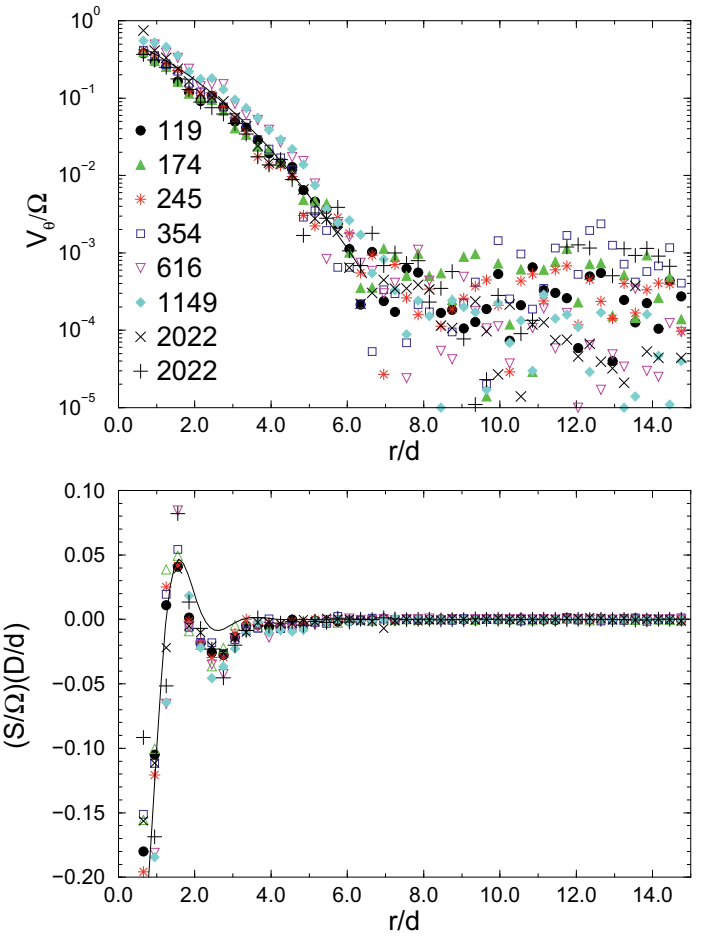

Fig. 17. Mean velocity and spin vs. radial distance r from the shearing wheel at $\gamma = 0.788$ for various values of $\tau = 2\pi/\Omega$ (τ in seconds).

relatively large r, i.e. $r > 6d$. Indeed, careful observation shows that there can be rearrangements of the grains in the outer regions over sufficiently long times. This is demonstrated in Fig. 14, which shows data for γ vs. radial distance from the wheel before any shearing, after 10 and 220 wheel rotations. In this instance, we observed rearrangement of the outer region between 20 and 40 wheel rotations, i.e, for wheel displacements corresponding to between $\sim$ 2200 and $\sim$ 4400 disk diameters. We speculate that these rare rearrangements lead to changes in the dynamics of particles in the inner region.

The mean spin, S, vs. r, Fig 17, bottom, also shows interesting structure. In a similar spirit to the velocity data, we present S in the normalized form, $(S/\Omega)(d/R_i)$ so that rate independence will be manifest if the data overlap for

all Ω. Note that a grain that is rolling in place against the inner wheel without slipping will have a normalized value of $(S/\Omega)(d/R_i) = -1$. Like the data for V_θ, the data for S are nearly but not perfectly rate invariant. Grains adjacent to the wheel are, on average, rotating backwards, whereas those in the next layer out are rotating in the same sense as the shearing wheel. The fact that these rotations die out quickly with distance from the wheel is quite reasonable, since the alternation of rotation with distance from the wheel must compete with frictional drag and rotational frustration from neighboring grains that are the same distance from the inner wheel.

The distributions for V_θ and S, Fig. 18, contain important information about the nature of the shear layer, and we return to this point below. These data pertain to disks that are in the first layer contacting the wheel, and data for pentagonal particles are qualitatively similar. There is a large peak near zero for both V_θ and S, another broad peak for nonzero values, and some hints of a third weak peak in V_θ.

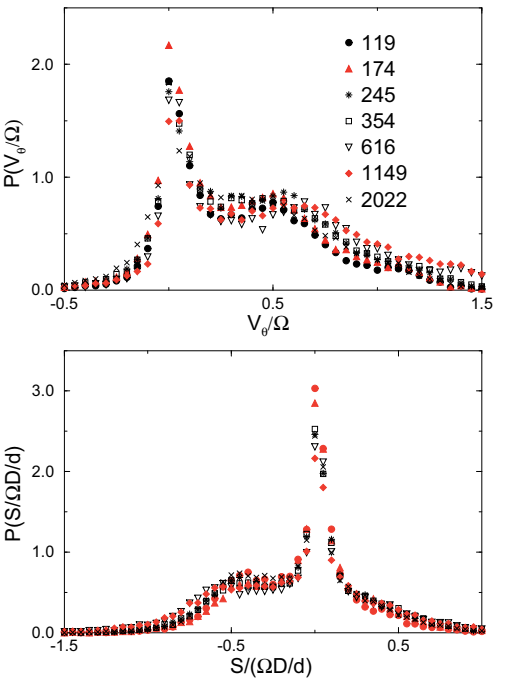

Fig. 18. Distributions for V_θ and S for disks within one disk diameter d of the shearing wheel. Different symbols correspond to different values of $\tau = 2\pi/\Omega$. Here, $\gamma = 0.788$.

Associated with the strong intermittency near γ_c is a dramatic variation of the kinematic properties with γ as seen in Fig 19 which gives the mean scaled velocity, V_θ/Ω vs. r. As $\gamma \to \gamma_c$ V_θ decreases dramatically, not because of slower movement of the particles when they actually move, but rather, because the disks move more sporadically.

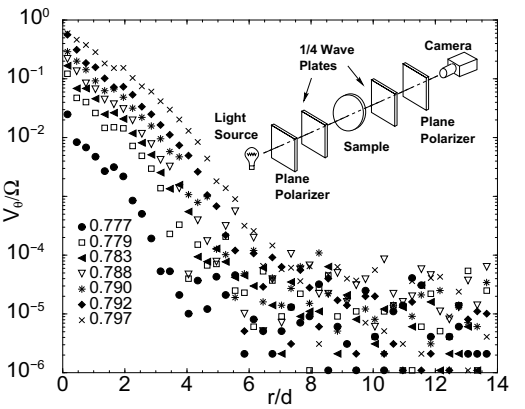

Fig. 19. Velocity vs. distance from the shearing wheel for various γ's. Inset: schematic of circular polarimeter used for the present photoelasticity measurements.

The slowing down as $\gamma \to \gamma_c$ satisfies a very clear scaling behavior, as shown in Fig. 20. Here, we have scaled the dimensionless velocity by $(\gamma - \gamma_c)/\gamma_c$, and the radial distance by a slowing varying function of r. The data for all the profiles collapse onto a universal scaling curve for both disks (A) and pentagons (B).

The velocity distribution next to the shearing wheel is also affected by changing γ, Fig 21 shows that the peak near zero velocity increases and the feature at 0.5 decreases as γ_c is approached from above. This means that as γ grows beyond γ_c, grains actually slip less, as captured in Fig. 22. This figure shows grey-scale representations of the two-variable distributions for velocity and spin, $P(V_\theta, S)$ for two different γ's for the disks that are next to the inner wheel. For $\gamma \simeq \gamma_c$ the preponderance of the weight is near the origin, whereas for a larger γ, much of the statistical weight has moved near the line

$$V_\theta/\Omega = 1 + S/(\Omega R_i/D) \tag{4}$$

corresponding to the condition that the grains are rolling without slipping against the inner wheel. This observation also gives some insight into why $\bar{V}_\theta/\Omega < 1$ at the shearing wheel (i.e. $\bar{\dot{S}} < 0$.

There are also dramatic changes in the statistical properties of the stress as γ varies, particularly as a result of the critical transition near γ_c. These properties surface in the mean stress, the stress distributions, and in the stress chain lengths.

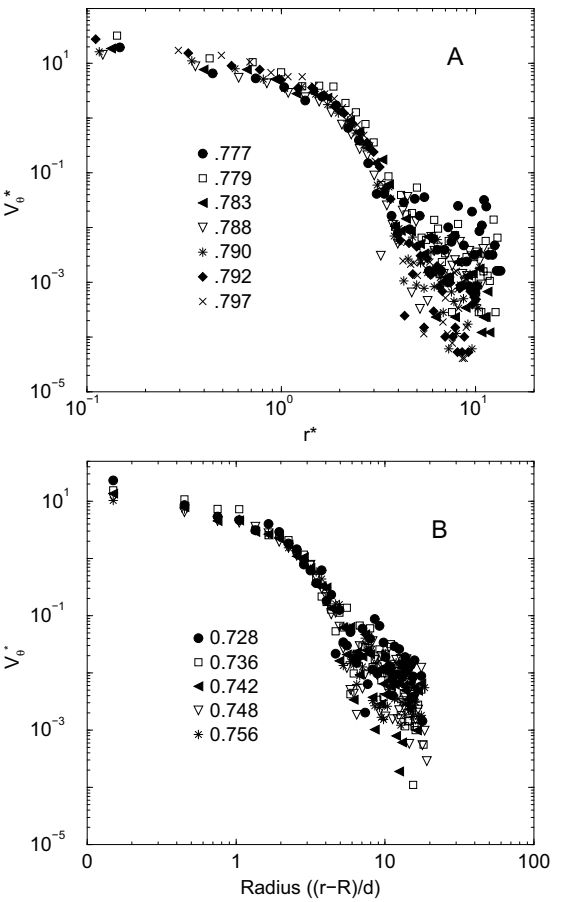

Fig. 20. Azimuthal velocity, scaled by the distance $(\gamma - \gamma_c)/\gamma_c$ from the critical transition, vs. a weakly recaled radial coordinate, $r/(1 + A(\gamma - \gamma_c)/\gamma_c)$. Excellent scaling is obtained. Top: data for disks; bottom: data for pentagons.

The mean stress, Fig. 23, rises very rapidly as γ exceeds γ_c. Here, we show data for both pentagons and disks, where we have averaged G^2 over about 1/10 of the sample, in a radial slice. In both cases, the mean stress varies as a power law, although the exponent, ~ 4.0 for disks, and ~ 2.2 for pentagons, is rather different between the two particle types.

The distribution of stresses for both types of particles varies significantly with γ. We characterize these distributions in two different ways. In Fig. 24 (left two panels), we show distributions on both log-log and log-lin scales of stress, σ, (force/length) obtained from G^2 integrated over the size of a disk. In the mid-range of γ's, these distributions are consistent with an exponential fall-off. At near-critical γ's, the distributions are neither power laws nor exponentials,

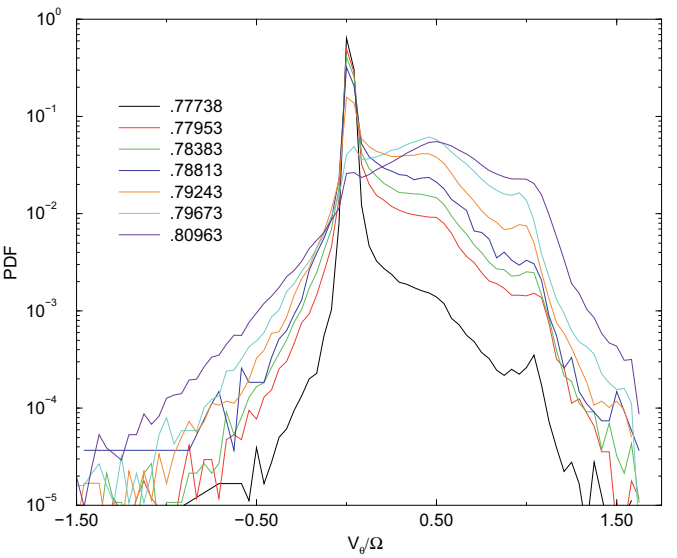

Fig. 21. Distribution of azimuthal velocity vs. velocity for various γs. These data are for disks.

but it possible that very close to γ_c, the distribution becomes a power law. At large stresses, the distributions begin to develop a rounded peak. The pentagon results are qualitatively similar. The second way in which we present the data is to give distributions for the force per particle obtained by averaging over images of ~ 200 particles at a time. These data correspond to the distribution that would be measured in 3D experiments where may particles contact a single force transducer, as in the measurements of Miller et al.[8]. We show this kind of distribution in the two right panels of Fig. 24. Here, the effects of averaging are clear, and the distributions are distinctly more gaussian looking. This is particularly true in the large force limit. Note that in the large-force regime, particles are beginning to deform enough that some of the load is distributed to particles that would otherwise have carried no load for smaller γ's.

One final statistical property that varies with γ is the length of stress chains, L, as shown in Fig. 25. Here, we show images for one case near γ_c and another case where γ is relatively high. Determining the length of a given chain is a complex task, and here we have taken an approach that consists of measuring the lengths of groups of particles that a) lie roughly along straight lines and b) have G^2 at or above the mean. Chains end at walls or when they split into two weaker chains. An optimum statistical measure might be a two-variable distribution of chain lengths for disks lying along straight lines and having force/particle within a given range. The determination of such a distribution from experimental data is, however, a formidable task, and we resort for the present to this simpler approach. The resulting distribution of chains lengths (inset–for disks) and the

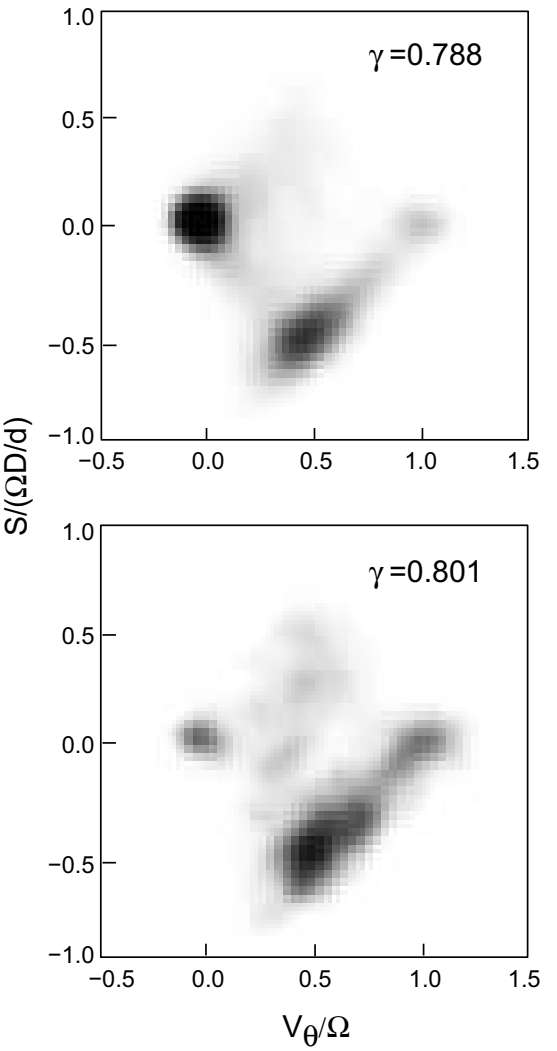

Fig. 22. Two-variable distribution, represented in greyscale, showing the change from slipping, where both the velocity and spin are zero, to rolling without slipping. The data are for disks with the indicated γ's.

mean chain lengths are given in Fig. 26. This figure shows that "chains" are typically very short, $\sim 2-3$ particles for larger γ's, but measurably longer at smaller γ's. The mean value of $L(\gamma)$ Fig. 26, bears out this perception. The mean length of a chain grows as $\gamma \to \gamma_c$. The exact functional form is not clearly determined at this time, but the data are consistent with a power law: $L \propto (\gamma - \gamma_c)^{-\alpha}$, where $\alpha \simeq 1/3$. Here, it is important to note that we are limited

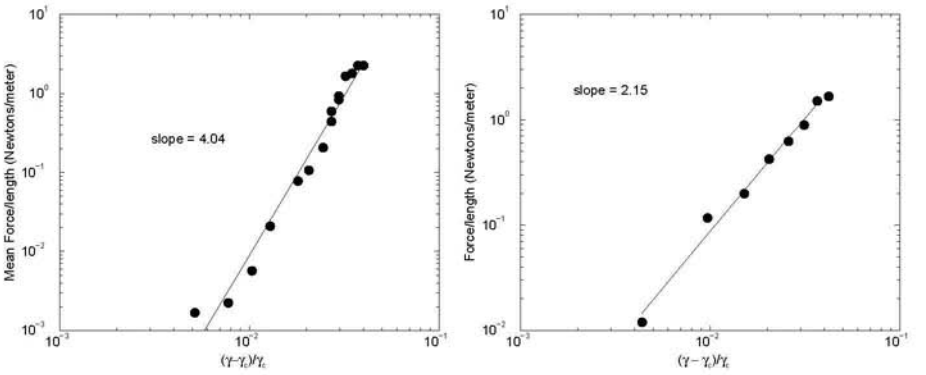

Fig. 23. Mean force vs. γ in 2D experiments. Left: data for disks; Right: data for pentagons.

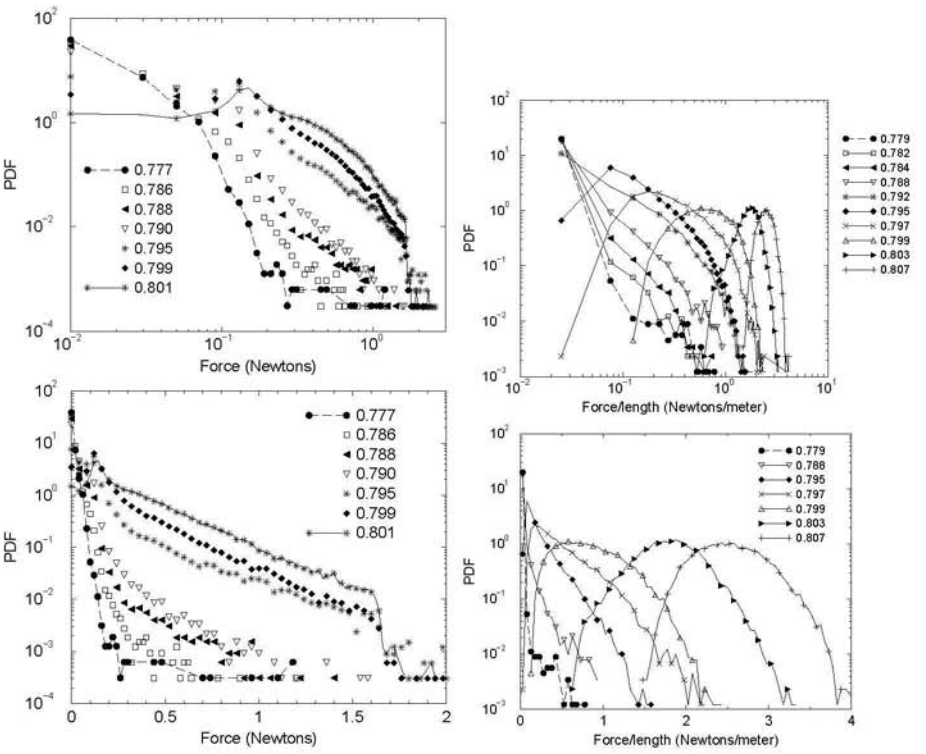

Fig. 24. Stress distributions on log-log and log-lin scales for the indicated γ's for disks. The left-hand data pertains to the force on a single disk. The right-hand data was obtained by averaging the force over roughly 260 disks.

by the finite size of the system. In future work, we plan to measure chain lengths for much larger systems.

To sum up briefly, the 2D Couette experiments exhibit a novel continuous transition for a granular material. This transition, which is intimately related to Reynolds dilatancy, is characterized by power laws for a number of properties, such as the mean stress, the velocity, etc. For other properties, such as the mean chain length, a power law is consistent with the data, even though another mathematical form would also be consistent with the results. Given this interesting critical-like behavior, it is interesting to speculate on its relation to other phenomena. One possibility is that this transition is related to rigidity percolation[55]. In this regard, we note recent numerical simulations by Aharonov and Sparks[54]. These authors report a rigidity phase transition for which some properties show discontinuous behavior as a function of density, even though the elastic constants undergo a second order transition.

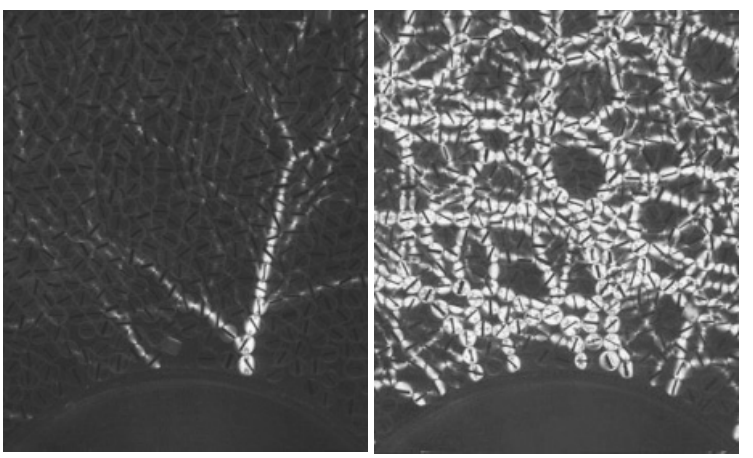

Fig. 25. Images showing stress chains at $\gamma = 0.777$, just above γ_c (left) and at high γ, $\gamma = 0.807$, right.

The power spectra associated with the force time series are also of interest since they give some insight into time scales in the system. The 2D spectra are remarkably like their 3D counterparts, with the exception of the near-critical data (not available for 3D). Specific results are presented elsewhere[46,47].

4 Transition to the Fluid State under Shaking

An alternative setting for understanding the transitions between different granular states can be found in shaken systems. A granular material makes the transition from solid to fluid under a variety of different kinds of vibration. Perhaps the most studied case is one of vertical vibrations. Here, the material will flow when

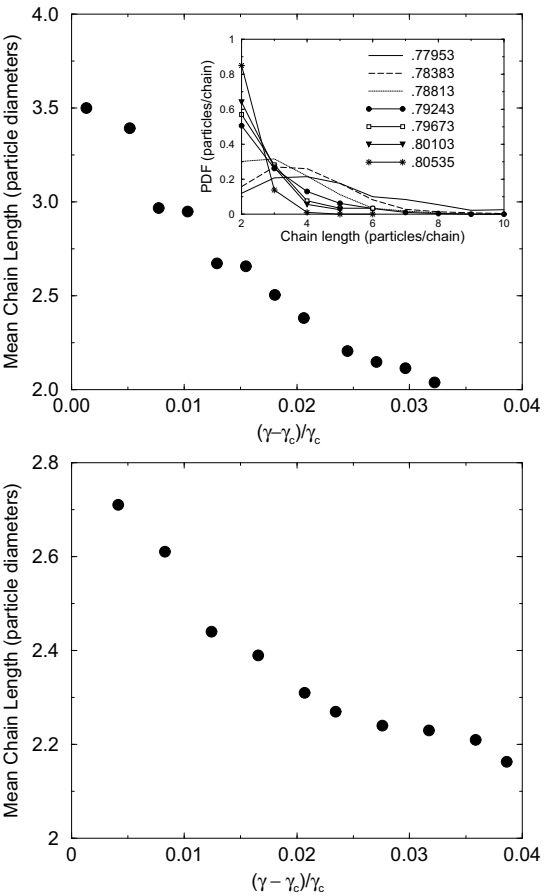

Fig. 26. a) Mean stress chain lengths: top–for disks, and bottom–for pentagons. Inset: typical distributions for the indicated γ's, in this case for disks.

it is subjected to accelerations that exceed the gravitational acceleration g. We have recently pursued a different approach by applying either purely horizontal vibrations or a novel combination of vertical and horizontal vibration[48,49] to a layer of granular material. The goal of these experiments is to specifically investigate the transitions back and forth between the solid and fluid phases. We will focus here primarily on the case of horizontal shaking.

In these experiments, a container of granular material subject to continuous horizontal and/or vertical displacements

$$s_i = A_i \cos(\omega_i t + \phi_i). \tag{5}$$

For each component of shaking there is a relevant control parameter

$$\Gamma_i = A_i {\omega_i}^2/g, \tag{6}$$

where g is the acceleration of gravity, and $i = v, h$ for vertical or horizontal. If $\omega_v = \omega_h$, the relative phase, $\phi_h - \phi_v$ is also important; otherwise, $\Delta\omega = \omega_v - \omega_h$ is important. Other parameters, such as $E = (A\omega)^2/gd$, which are relevant in describing higher order phenomena in vertically shaken materials (e.g. traveling waves[50], coarsening[51]) are not necessary to describe the onset of flow in these experiments. Typically, the layer had a height $h/d \simeq 40$; as long as $h/d > 20$, we found no dependence on h.

Our experimental setup is shown in Fig 27. The heart of the experiment is a rectangular Plexiglas cell with cross-sectional dimensions of 1.93 cm by 12.1 cm. The base of the cell is made of a porous medium through which gas can flow in order to fluidize the granular medium. This provides an independent control over the dilation of the material. The cell is mounted on a Plexiglas base of the same cross-sectional dimensions as the cell, which is in turn mounted on a movable table; the base acts as the gas distributor to the system. The table is mounted on four linear bearings sliding on corresponding horizontal shafts. An independent electro-mechanical actuator provides a sinusoidal drive for each degree of freedom.

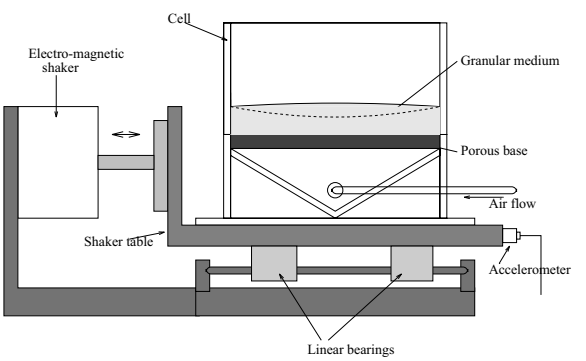

Fig. 27. Schematic of the apparatus for horizontal shaking.

For these experiments, we used several types of granular materials, of roughly uniform size, including spherical glass beads, smooth nearly spherical Ottawa sand, and sieved rough sand. These particles had diameters ranging from $0.2mm \leq d \leq 1.0mm$. We typically prepared the system by pouring in the material, leveling the sample, and then compacting it at modest Γ. The material did not form a perfectly ordered packing because the particles where not sufficiently spherical or monodisperse.

In order to characterize the flow, we used both high-speed (250 frames/s) and conventional video. Much of the data discussed here were obtained by observing the flow near the walls through the transparent Plexiglas container with a small surveillance CCD camera mounted on the shaker platform. Using video images, a framegrabber and computer software for particle tracking, we extracted information such as heap heights, shapes of the top surface, flow speeds, and mean

flow fields. In a typical run for pure horizontal shaking, we observed the evolution of the system as A_h was increased from zero while keeping ω_h fixed. Once flow was well initiated, we then decreased A_h.

For simple horizontal shaking, we are particularly interested in the initial instability to fluidization. The transition to this flow state is hysteretic in Γ_h, and leads to a characteristic flow field that consists of sloshing of the grains in the direction of the shaking, plus slower convective flow, including flow in the direction transverse to the shaking. Fig 28 shows a sketch of the mean convective flow lines (i.e. discounting the sloshing flow) observed from the top and side for Γ_h somewhat above onset in a cell that is 30 grains wide and 50 grains deep. Grains rise up in the middle of the cell and flow along the surface towards the side walls and then sink at the wall boundaries. The top surface of the liquefied layer has a dome shape that is concave down, and the bottom surface of this layer is concave upwards near the wall boundaries. Thus, the thickness of the fluidized layer is largest in the middle of the cell and smallest at the end walls, as seen previously by Evesque[52].

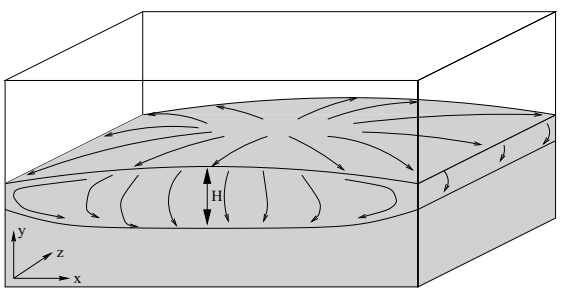

Fig. 28. Sketch of convection flow lines in the liquified layer induced by horizontal shaking. Grains rise in the middle of the cell, flow along the surface towards the side walls, and then sink at the wall boundaries.

A useful measure of the strength of the flow is the thickness, H, of the liquefied material in the middle of the cell. Fig. 29 shows the typical hysteretic behavior for H as a function of Γ_h. With increasing Γ_h, there is a well defined transition to finite amplitude (i.e. $H \neq 0$) flow at Γ_{cu}. With Γ_h decreasing below Γ_{cu} once flow has begun, the thickness of the layer also decreases but H not vanish until Γ_h reaches a lower critical value Γ_{cd} where the relative motion completely stops. As Γ_h is decreased from above towards Γ_{cd}, grains near the end walls stop moving first, while grains in the middle keep moving. Some hysteresis was present in the 2D studies by Ristow et al.[53], but the hysteresis in the our 3D experiments is larger, particularly for rougher materials. We also note that below the onset of sloshing there exist a small number of free particles that slide about on the top of the layer (we refer to these as 'sliders'). Roughly, these particles act like a 'gas' in the sense that they have a low density. Yet, they seem to play a surprising role in the solid-to-fluid transition, as noted below.

The transition to the sloshing state is described within the context of Mohr-Coulomb failure, although this model does not tell us the depth of the sloshing layer. The reverse transition in which the sloshing layer vanishes is not well described by any theory, to our knowledge. In particular, if we use the conventional concept of S-o-S friction (i.e. if we imagined that the sloshing layer were like a solid block that interacted with the rest of the material by S-o-S friction) in which there is an instantaneous change from static to kinetic friction or vis-versa, then we would not expect to observe hysteresis. The ingredient that is missing from this block-model picture is the need for the layer to dilate when changing to fluid and its tendency to compact when it returns to the solid state.

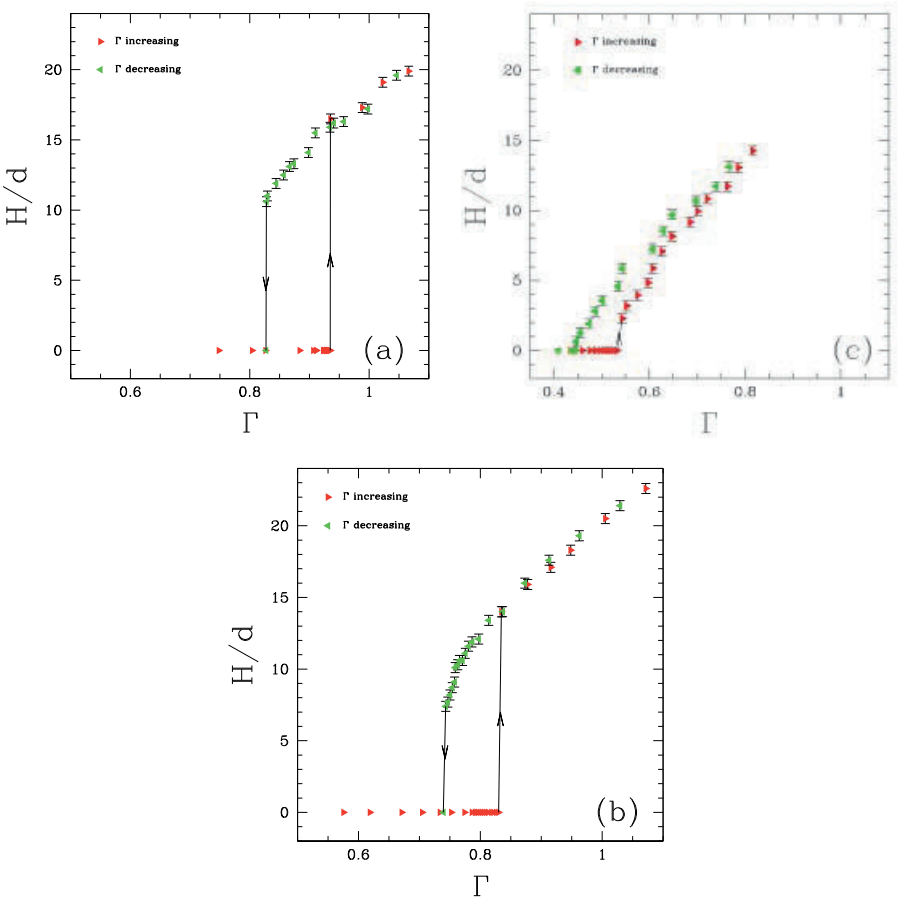

Fig. 29. Bifurcation diagrams for the initial sloshing transition of horizontally agitated dry granular materials: (a) rough sand, (b) glass beads, and (c) smooth Ottowa sand. The transition points and size of the hysteresis loop tend to increase with increasing material roughness.

The location of Γ_{cu} and Γ_{cd} are reproducible for a particular material (within the accessible range of these experiments). That is, different ω's or A's yield the same critical Γ_h's for a given material and fill height. The values of Γ_{cu} and Γ_{cd} are also not affected by the total height of the layer, if that quantity is roughly 20 grains or higher. However, these transitions values of Γ_h do depend on the physical properties of the material, as shown in Fig 29 (Left). For instance, Γ_{cu} and Γ_{cd} increase as the roughness of the granular materials increases. This can be attributed to the greater interlocking of rough grains and because rough grains likely have higher S-o-S friction coefficients. It is also possible to map out the full bifurcation diagram, including the unstable states, as shown in Fig. 30. Here, the technique that we used involved applying known perturbations to the system, and then observing to see whether the fluid or solid was the end state following the perturbation.

To obtain additional insight into the relative importance of dilatancy and friction, we have carried out two additional experiments. In the first of these, we fluidized the granular bed by passing air through it, using the flow-controlled air supply (see Fig. 27), where the porous base of the cell acts as the gas distributor. Fig 31, which presents Γ_{cu} and Γ_{cd} as functions of the air flow through the bed, shows a very strong dependence of these quantities on the air flow. In particular, a modest air flow reduces the critical Γ's and ultimately effectively removes the hysteresis in the initial transition. Here, ΔP is the pressure difference across the granular layer, and W is its weight per area. We note that the measured dilation of the bed due to the air flow is small: the maximum dilation for these experiments corresponds to less than one granular layer for a bed of 45 layers, i.e. less than 2% dilation. In a second experiment to probe the effects of dilatancy, we placed a small amount of overburden (here a thin piece of plastic of weight comparable to about half a layer of grains). This small strip of plastic suppressed the motion of the sliders. Surprisingly, we found that the hysteresis was lifted, and that the transition to flow occured at a slightly reduced value of Γ_h relative to Γ_{cd}, Fig. 30 (Right). This result is particularly striking because we might expect that the addition of weight would delay the transition, rather than enhance it.

An additional point concerns nucleation times. These include the time to nucleate the fluid layer from the solid when Γ_h is stepped slightly past Γ_{cu}, and the time for all of the fluid phase to decay away if Γ_h is stepped slightly below Γ_{hd}. We show representative results for both processes in Fig. 32. Again, there is a surprise: there is a very well defined power law for each process, with the relaxation time always varying as $\tau = A/\epsilon$, where $\epsilon = (\Gamma_h - \Gamma_{ci})/\Gamma_{ci}$, $(i = u, d)$. This set of results is striking because it suggests a universal behavior. An interesting and unresolved question is the physics behind the exponent minus one.

An important issue concerns the driving mechanism or 'motors' for both the convective flow parallel to the shaker plane and the convective flow perpendicular to the shaker plane. The origin of the sloshing motion is relatively obvious: during each half of the shaker cycle grains pile up on one side wall and open a gap near the opposite side wall, such that grains near the opening avalanche down to

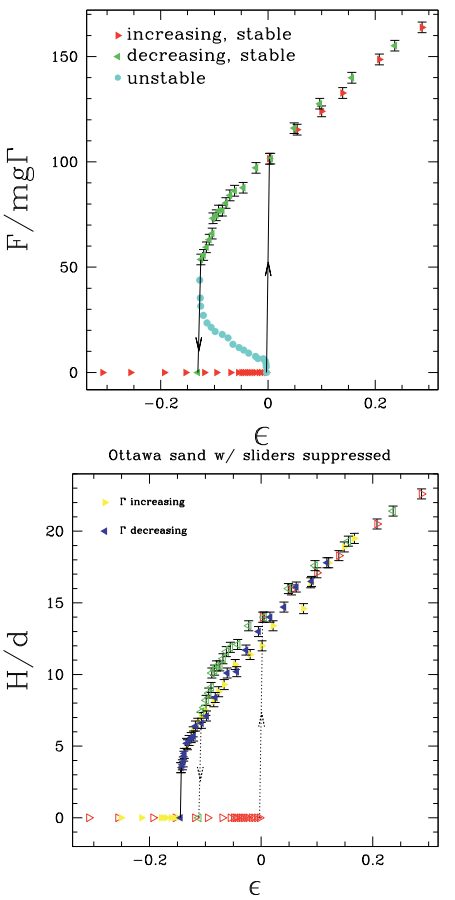

Fig. 30. Left: Full bifurcation diagram, including the unstable (negative slope) branch, for Ottawa sand vs. Γ_h. Right: corresponding diagram in which the surface motion, i.e. the sliders, is suppressed. Here, the data for the hysteretic transition (note arrows) are ploted with the nonhysteretic results that are obtained when here is a slight overburden.

fill part of the gap. This process repeats at one endwall every half cycle. The convective flow transverse to the shaker plane is more subtle and is due to a completely different process. To investigate this cross flow, we have carried out additional experiments in which we only oscillate one of the long side walls. We find that that the oscillation of the sidewall at essentially any Γ or A leads to a convective flow where grains fall at the sidewall, and are then pushed inward, and upward into the interior of the layer. We find that the average speed at which grains fall along the shearing wall, v_z, is simply a linear function of the speed, $A\omega$ of the wall relative to the grains, and these results are shown in Fig. 33. We have also carried out MD simulations[49] that reinforce this interpretation. This

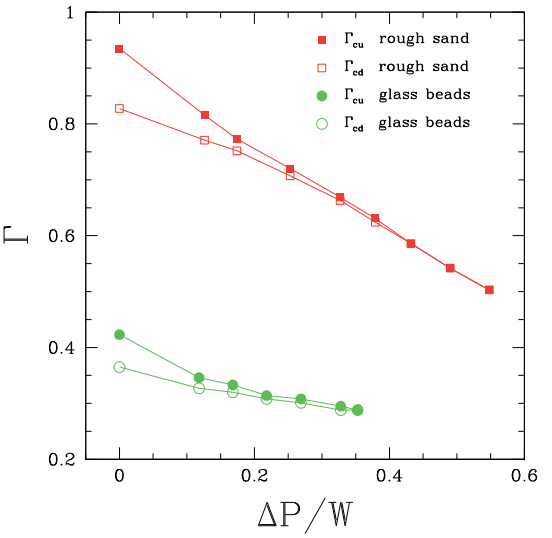

Fig. 31. Bifurcation points Γ_{cu} and Γ_{cd} versus dimensionless lifting pressure and air flow rate for rough sand and for glass beads.

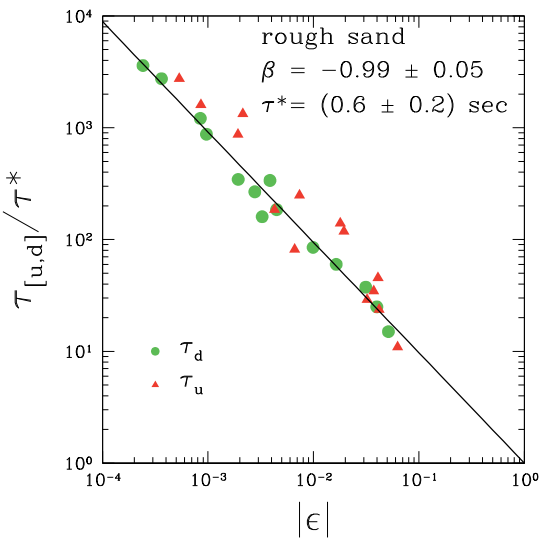

Fig. 32. Relaxation times for rough sand vs. $\epsilon = (\Gamma_{ci} - \Gamma_h)/\Gamma_{ci}$ for the nucleation of the fluidized layer with increasing Γ_h and the vanishing of the fluidized layer with decreasing Γ_h.

kind of shear-induced convection is likely to occur generally when a 3D material is sheared in the presence of a gravitational field. We compare the results of these calculations with experiment in Fig. 33

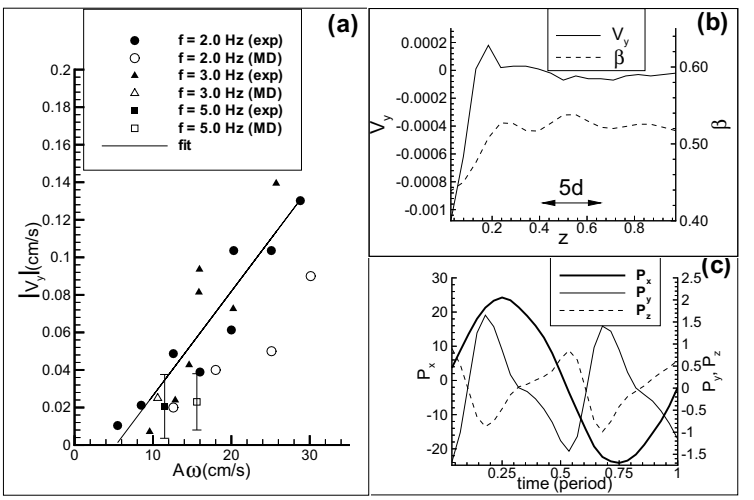

Fig. 33. Comparison of MD calculations and experiment. The left part of the figure shows MD results and experimental results for the downflow of particles along the vertical wall that is induced by shearing. The right part of the figures shosw typical velocity profiles obtained from the simulations

When the whole layer of material is now shaken, the following scenario describes the onset of the instability and the mechanism for the cross-convective flow. The flow begins when the grains overcome frictional and dilatancy effects. This flow is nucleated at a single point near the solid surface. With the formation of the fluid layer, there is a shear flow of surface grains relative to the long walls (i.e. the direction of shaking) which is a maximum in the longitudinal middle of the cell. Near the side walls, there is a strong dilatancy of the moving grains, due to shearing. The highly dilated grains near the side wall are mobile, and grains reaching this boundary at the top surface percolate downward. However, grains falling along these walls eventually are pushed inwards.

5 Conclusions

We have explored aspects of granular materials that touch on the solid and dense fluid phases. Necessarily, we have omitted many interesting results that concern the gas-like phase. Conventional theories have typically considered only the mean behavior of granular systems. Recent work has focused on the microscopic structure and statistics of these materials. From this point, the mean behavior is obtained in the long wavelength limit. Several models and techniques addressing

the microscopic structure of granular materials include the q-model (and related constructions), the OSL model, and contact dynamics. In general, these models, as well as experiments, agree that the distribution of forces on individual grains should be described at large forces by an exponential: $P \propto \exp(-F/F_o)$. What is still unclear theorectially is the nature of correlations, their connection to force chains and how to fold history into the description. Such issues necessarily affect the long wavelength limit of any description. As a consequence, there are now competing theories that propose an array of very different partial differential equations for stress fields within a granular material.

We have explored these theories through experiments that measure forces at the base of 3D sand piles, forces and structural information in 2D heaps, and Green's function measurements for 2D heaps. The first experiments confirm the importance of history dependence, and they are consistent with the OSL model. The 2D heap experiments show that structure even more than force profiles is the quantity most strongly affected by history. Finally, the Green's function measurements, which to date must still be considered preliminary, suggest that the mean behavior of a granular system is described by a wave-like process with dispersion.

We have also carried out a series of experiments in 3D and in 2D that have laid out the nature of fluctuations in slowly sheared systems. In 3D, there is an interesting rate-independence for the fluctuation forces; that is, the shear rate provides the characteristic time scale for the system, even at the level of the flucutations. Typically, we see force distributions that are consistent with an exponential distribution. But we do not observe the narrowing of the distributions with increasing number of particles that would be consistent with a total lack of correlation between grains. From the perspective of these experiments, the issues of correlations is not resolved. In 2D, we also observe exponential distributions for a range of densities. However, for densities close to the critical value where stresses are just maintained in the system, and for highly compressed samples, the distributions differ from exponential. The near-critical regime is particularly interesting. Here, a number of quantities, including the compressibility, the characteristic length of a stress chain, the distribution of avalanches, and the mean velocity appear to satisfy powerlaws–in analogy to what one might expect for a fluid near the liquid-gas critical point. This transition is important for dense granular materials because it provides the mechanism for shear deformation. As such, it is connected to Reynolds dilatancy, shear bands and similar phenomena.

Finally, we have explored the transition between fluid and solid phases through the use of horizontal and combined horizontal and vertical shaking. Here, we have considered only results for pure horizontal shaking, where we see a hysteretic transition between the fluid and solid states. Interestingly, there is power law scaling for the time to nucleate either phase, and the hysteresis is lifted by either dilating the material slightly or by suppressing the layer of sliding particles on the surface. Suppression of the hysteresis with dilation, and the divergent time scales implies that the dilation process is key to the transition between these phases. It is this process that separtates Coulomb failure of granular

materials from ordinary solid-on-solid frictional failure. The specific and highly reproducible forms for the nucleation times suggests a global picture should apply to both directions of the phase change, i.e. solid-to-fluid, and fluid-to-solid. Conventional models of Mohr-Coulomb failure treat only the solid-to-fluid transition, and even then do not account for the details of the transition.

Acknowledgments

The work of DH and RPB was supported by the National Science Foundation under Grant DMR-9802602, and DMS-9803305, and by NASA under Grant NAG3-1917 and NAG3-2372. The work of E. Clement, G. Reydellet and L. Vanel was supported under PICS-CNRS n^o 563.

*Current address: Argonne National Laboratory, Materials Science Division, 9700 Cass AVe. Argonne, IL 60439, USA

**Current address: Dept. of Physics, UCLA, 405 Hilgard Ave. Los Angeles, CA 90095-1547

†Current address: Danish Research Centre for Magnetic Resonance, Hvidovre Hospital, Dept. 340, Kettegaard Alle 30, DK 2650 Hvigovre, Denmark

References

1. For broad reviews, overviews, and proceedings see: Focus issue of *Chaos*–R. Behringer, H. Jaeger, and S. Nagel, Chaos **9**, 509 (1999); L. P. Kadanof, Rev. Mod. Phys. **71**, 435 (1999); *Physics of Dry Granular Media*, H. J. Herrmann, H.-P. Hovi, and S. Luding, eds. Kluwer, 1998; Proceedings of Powders and Grains, R. P. Behringer and J. T. Jenkins, eds. Balkema (1997); H. M. Jaeger and S. R. Nagel and R. P. Behringer, *Physics Today*, **49**, 32 (1996); Rev. Mod. Phys. **68** 1259 (1996); *Physics of Granular Media*, D. Bideau and J. Dodds, eds. Les Houches Series, Nova (1991); *Granular Matter: An Interdisciplinary Approach*, A. Mehta, Ed. Springer, NY (1994); R. P. Behringer, Nonlinear Science Today, **3**, 1 (1993).
2. B. J. Ennnis, J. Green, and R. Davies, Chem. Eng.Prog. **69**, 44 (1994).
3. J. Eibl, The Structural Engineer, **62A**, 169 (1984).
4. A sampling of relevant works includes: S. Ogawa in *Proceedings of US-Japan Seminar on Continuum-Mechanical and Statistical Approaches in the Mechanics of Granular Materials*, ed. S. C. Cowin and M. Satake (Gakujutsu Bunker Fukyukai, Tokyo, Japan) 1978; P. K. Haff, J. Fluid Mech. **134**, 401 (1983); S. B. Savage, Adv. Appl. Mech. **24**, 289 (1984).
5. W. Losert, D.G. W. Cooper , J. Delour, A. Kudrolli, and J. P. Gollub, *Chaos* **9**, 682 (1999); M. Shattuck, H. Swinney, private communication.
6. I. Goldhirsch and G. Zanetti, Phys. Rev. Lett. **70**, 1619 (1993).
7. C.-h.Liu, S. R. Nagel, D. A. Schecter, S. N. Coppersmith, S. Majumdar, O. Narayan and T. A. Witten, Science **269**, 513 (1995), Phys. Rev. E **53**, 4673 (1996).
8. B. J. Miller, C. O'Hern, and R. P. Behringer, Phys. Rev. Lett. (1996).
9. D. K. Mueth H. M. Jaeger, S. R. Nagel, Phys. Rev. E **57**, 3164 (1998).
10. A. Ngadi and J. Rajchenbach, Phys. Rev. Lett. **80**, 273 (1998).
11. D. Howell, C. Veje and R. Behringer, Phys. Rev. Lett. **82**, 5241 (1999).

12. I. Albert, P. Tegzes, B. Kahng, R. Albert, J. G. Sample, M. Pfeifer, A.-L. Barabasi, T. Vicsek and P. Schiffer, Phys. Rev. Lett. **84** 5122 2000; Phys. Rev. Lett **82** 205 (1999).
13. A. V. Tkachenko and T. A. Witten, to appear.
14. P. Dantu, Géotechnique, **18**, 50 (1968); A. Drescher and G. De Josselin De Jong, J.Mech. Phys. Solids, **20**, 337 (1972); T. Travers et al. J. Phys. A. **19**, L1033.
15. R. Jackson, in *Theory of Dispersed Multiphase Flow, Ed. Re. E. Meyer* (Academic Press, NY, 1983, p. 291
16. G. Amontons, Mémoires de l'Academie Royale, A **12**, 257 (1699).
17. R. M. Nedderman, *Statics and Kinematics of Granular Materials*, Cambridge, 1992.
18. For a review see R. Jackson, In *Theory of Dispersed Multiphase Flow* ed. R. Meyer, Academic Press (1983).
19. See C. Caroli and P. Nozières and T. Baumberger in the Proceedings of the Workshop on Friction, Arching, Contact Dynamics, D. E. Wolf and P. Grassberger, eds. World Scientific, 1997.
20. S. Ouaguenouni and J.-N. Roux, Europhys. Lett. **32**, 449 (1995) and in the Proceedings of the Workshop on Friction, Arching, Contact Dynamics, D. E. Wolf and P. Grassberger, eds. World Scientific, 1997; A. V. Tkachenko and T. A. Witten, Phys. Rev. E **60** 687 (1999).
21. A. Liu and S. Nagel, Nature **396**, 21 (1998).
22. Christophe Eloy and Eric Clément, J. Phys. I France **7**, 1541 (1997).
23. Joshua Socolar, Phys. Rev. E **57**, 3204 (1998).
24. Mario Nicodemi, Phys. Rev. Lett. **80**, 1340 (1998).
25. F. Radjai, M. Jean, J.-J. Moreau, and S. Roux, Phys. Rev. Lett. **77**, 274 (1996); R. Radjai, D. Wolf, M. Jean and J.-J. Moreau, in Friction, Arching and Contact Dynamics, ed. D. E. Wlof and P. Grassberger, World Scientific (1997).
26. J. P. Wittmer, P. Claudin, M. E. Cates, and J. -B. Bouchaud, Nature **382**, 336 (1996); Phys. Rev. E **57**, 4441 (1997).
27. V. M. Kenkre, J. E. Scott, E. A. Pease, and A. J. Hurd, Phys. Rev. E **57** 5841 and 5850 (1998).
28. S. B. Savage, in *Physics of Dry Granular Media*, H. J. Herrmann, H.-P. Hovi, and S. Luding, eds. Kluwer, 1998.
29. L. Vanel, D. Howell, D. Clark, R. Behringer, and E. Clément, Phys. Rev. E. **60**, R5040 (1999).
30. R. Brockbank, J. M. Huntley, and R. Ball, J. Physique I France **10**, 1521 (1997).
31. R. B. Heywood, *Designing by Photoelasticity*, Chapman and Hall Ltd, London, 1952.
32. S. G. Paikowsky and K. J. DiRocco, Conference on Digital Image Processing: Techniques and Applications in Civil Engineering, Kona, Hawaii, February 28- March 5, 1993, ASCE publication 236 (1993).
33. G. W. Baxter and R. P. Behringer, Eur. J. Mech. B, **10**, 181 (1991). G. W. Baxter, R. Leone and R. P. Behringer, Europhys. Lett. **21**, 569 (1993).
34. C. H. Liu and S. R. Nagel, Phys. Rev. Lett. **68**, 3201 (1992).
35. X. Jia, C. Caroli and B. Velicky, Phys. Rev. Lett. **82** 1863 (1999).
36. H. J. Herrmann, S. Luding, J. P. Hovi, eds. *Dry Granular Media*, Nato ASI series, Kluver, Amsterdam (1998).
37. D. Howell et al. in *Proceedings of the Workshop on Friction, Arching and Dilatancy* ed. E. E. Wolf and P. Grassberger, World Scientific, 1997; C. T. Veje, et al. in *Dry Granular Media*, H. J. Herrmann, S. Luding, J. P. Hovi, eds. , Nato ASI series, Kluver, Amsterdam (1998).

38. D. Howell, Ph.D. Thesis, Duke University, 1999, unpublished.
39. R. M. Nedderman and C. Laohakul, Powder Technol., **25**, 91 (1980).
40. C. T. Veje, Daniel W. Howell and R. P. Behringer, Phys. Rev. E **59**, 739 (1999).
41. S. Schöllmann, S. Luding and H. J. Herrmann, to be published.
42. C. Veje, D. Howell, R. Behringer, S. Cshöllmann, S. Luding and H. Herrmann, in *Physics of Dry Granular Media*, H. J. Herrmann, H.-P. Hovi, and S. Luding, eds. Kluwer, 1998.
43. H. Buggisch and G. Löffelmann, Chem. Eng. Processing (1989).
44. E. Ben-Naim, J. B. Knight and E. R. Nowak, J. Chem. Phys., **100**, 6778 (1996); Edmund R. Nowak et al., Phys. Rev. E **57**, 1971 (1998).
45. H. Jaeger, private communication
46. D. Howell, R. P. Behringer, and C. T. Veje, Chaos **9**, 559 (1999).
47. R. P. Behringer, D. Howell, L. Kondic, S. Tennakoon, and C. Veje, Physica D **133**, 1 (1999).
48. S. G.K. Tennakoon and R. P. Behringer, Phys. Rev. Lett. **81**, 794 (1998).
49. Europhys. Lett. S. G.K. Tennakoon, L. Kondic, and R. P. Behringer**45**, 470 (1999).
50. H. K. Pak and R. P. Behringer, Phys. Rev. Lett. **71**, 1832 (1993).
51. E. van Doorn and R. P. Behringer, Phys. Lett. A **235**, 469 (1997).
52. P. Evesque, *Contemp. Phys.* **33**, 245 (1992).
53. G. H. Ristow, G. Strassburger, and I. Rehberg, Phys. Rev. Lett. **79**, 833 (1997).
54. E. Aharonov and D. Sparks, Phys. Rev. E **60**, 6890 (1999).
55. S. B. Obuknov, Phys. Rev. Lett. **74**, 4472 (1995); C. Moukarzel, P. M. Duxbury, and P. L. Leath, Phys. Rev. Lett. **78**, 1480 (1997).

Cluster Instability in Freely Evolving Granular Gases

J. Javier Brey[1]

Física Teórica. Universidad de Sevilla. Apartado de Correos 1065. 41080 Sevilla, Spain

1 Introduction

A granular medium is formed by a large number of macroscopic particles or *grains*. Here large must be understood at a macroscopic level, i.e. a few thousands or even a few hundreds is already a large number in the present context, as compared with molecular systems which contain a number of atoms or molecules of the order of the Avogadro number. In this lecture, we will restrict ourselves to dry granular systems in which there is not any other fluid around the grains. Also, electrical effects are not considered. Under these circumstances, the grain-grain interaction can be taken as purely repulsive with no attractive part.

Depending on the external conditions being applied, granular systems can exhibit a wide range of quite peculiar behaviours [1]. In particular, in the so-called rapid granular flow regime, grains behave similarly to molecules of an ordinary classical fluid [2,3], in the sense that they move freely between collisions, that in first approximation can be considered as binary and instantaneous. Nevertheless, there is one crucial difference between molecular fluids and rapid granular flows. While molecular collisions are elastic, collisions between grains are inelastic, i.e. kinetic energy is not conserved. Moreover, this inelasticity is inherent to granular fluids, in the sense that it is responsible for most of their peculiar properties.

Many experiments and also Molecular Dynamics simulations have suggested quite strongly that there is a hydrodynamic-like level of description for rapid granular flows, at least for not very high inelasticity. Nevertheless, the explicit form of the hydrodynamic-like equations must account for the essential features of a system composed by grains. In particular, it is clear that the balance equation for the energy needs to incorporate a sink term describing energy loss in collisions. The explicit form of this term and how the expressions for the energy and momentum fluxes must be modified, can not be obtained by a direct extrapolation of the Navier-Stokes equations for a molecular fluid.

The theoretical basis for molecular hydrodynamics starting from a microscopic level is provided by kinetic theory. The possibility of using the same methods to study granular fluids appears in a natural way. The first systematic formulation of this approach was made by Haff [2], whose ideas were further elaborated in [4] and [5]. Nowadays, the kinetic theory for systems of inelastic particles seems well established [6], and even a general nonequilibrium statistical mechanics based on a pseudo-Liouville equation has been formulated recently [7].

Inelasticity in collisions makes the problem of solving the kinetic equations much more involved mathematically. In fact, no exact solution of the Boltzmann

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 392–407, 2001.
© Springer-Verlag Berlin Heidelberg 2001

equation for a granular medium is explicitly known even for the simplest homogeneous situation. A modified Chapmann-Enskog expansion has been used to derive hydrodynamic equations from the inelastic Boltzmann equation [8,9], but it is not a priori clear that their range of applicability be as general as in the case of ordinary fluids. The cooling rate following from energy dissipation in collisions introduces a new time scale that is faster than the usual hydrodynamic scale. If there were no separation between this new scale and the purely kinetic contributions to the dynamics of the system, the reduction of the description to the hydrodynamic fields would not be permissible. Although a definitive answer to this question has not been given yet, hydrodynamics has proven to be very useful to understand and explain the behaviour of fast granular flows in many different situations and for a wide range of inelasticities.

In this paper we will discuss one far from trivial phenomenon shown by granular flows and that can be understood starting from the hydrodynamic description of the system. When a granular fluid is allowed to evolve freely, it tends to form density clusters [10,11]. The physical origin of this instability can be explained as follows. Suppose that there is a localized density fluctuation in a given region of the fluid and the density increases. Then, the number of collisions per unit of time also increases, and the energy decays faster than in the surroundings due to inelastic dissipation in collisions. If it happens that the increase of the density is not enough to compensate the temperature decay, the pressure in that region falls bellow the average value, a flux of particles towards the region is produced, and the clusters begin to develop. Nevertheless, this appealing physical argument is not easy to translate into a quantitative language. In fact, two different detailed hydrodynamic sequences have been proposed as responsible for the clustering development in the initial stages. Discriminating between them is one of the aims of this paper.

The presentation here will be based on the assumption that the Boltzmann equation, properly modified to account for the inelasticity in collisions, provides an accurate description of a low density gas of inelastic hard spheres or disks. From the Boltzmann equation, hydrodynamic like equations have been derived by using an extended Chapman-Enskog procedure, based on the time scale separation between the hydrodynamic modes and the kinetic ones. On the other hand, numerical solutions of the Boltzmann equation for given specific situations can be obtained by using the Direct Simulation Monte Carlo method. The comparison of the results obtained by both methods provides a strong test for the validity of the hydrodynamic description, allowing to check the assumptions introduced.

2 Hydrodynamic Equations

The simplest model for rapid granular flows is a system of smooth inelastic hard disks ($d = 2$) or spheres ($d = 3$) of diameter σ and mass m. The collision law is defined by momentum conservation and by

$$\boldsymbol{g}' \cdot \widehat{\boldsymbol{\sigma}} = -\alpha \boldsymbol{g} \cdot \widehat{\boldsymbol{\sigma}}, \tag{1}$$

where $\boldsymbol{g}$ and $\boldsymbol{g}'$ are the pre- and postcollision relative velocities and $\widehat{\boldsymbol{\sigma}}$ is a unit vector along the line joining the centers of the two particles at contact, away from the target particle. The inelasticity of collisions is described by means of the coefficient of normal restitution α defined in the interval $(0, 1]$. The case $\alpha = 1$ represents elastic particles. Although for real grains α is a function of the normal component of the impact relative velocity [12,13], here it will be taken as a constant for simplicity reasons. The loss of energy in a collision is $\Delta E = -(1-\alpha^2)m(\boldsymbol{g}\cdot\widehat{\boldsymbol{\sigma}})^2/4$. Of course, it vanishes in the limit $\alpha \to 1$.

Let us now consider a dilute granular gas. Once the collision rules are defined, it is possible to apply the same empirical reasoning as for molecular systems and derive the extension of the Boltzmann equation for inelastic particles [14]. Also, a pseudo-Liouville equation for the N-particle distribution function of the system has been derived. Integration of this equation yields the BBGKY hierarchy for the reduced distribution function [7]. In the low density limit, and introducing the molecular chaos assumption, the first equation of the hierarchy reduces to the inelastic Boltzmann equation for the one particle distribution function of the granular gas $f(\boldsymbol{r}, \boldsymbol{v}, t)$,

$$\left(\frac{\partial}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\right) f(\boldsymbol{r}, \boldsymbol{v}, t) = J[\boldsymbol{r}, \boldsymbol{v}|f(t)]. \tag{2}$$

Here $J[\boldsymbol{r}, \boldsymbol{v}|f]$ is the Boltzmann collision operator for inelastic hard disks or spheres [7]. Defining in the standard way the local number density $n(\boldsymbol{r}, t)$, the local flow velocity $\boldsymbol{u}(\boldsymbol{r}, t)$, and the local temperature $T(\boldsymbol{r}, t)$, balance equations for these quantities are obtained by taking moments of the Boltzmann equation with respect to 1, $\boldsymbol{v}$, and v^2 [4,5],

$$\frac{\partial n}{\partial t} + \boldsymbol{\nabla}\cdot(n\boldsymbol{u}) = 0, \tag{3}$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} + (mn)^{-1}\boldsymbol{\nabla}\cdot\mathsf{P} = 0, \tag{4}$$

$$\frac{\partial T}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}T + \frac{2}{dn}(\mathsf{P}:\boldsymbol{\nabla}\boldsymbol{u} + \boldsymbol{\nabla}\cdot\boldsymbol{q}) + T\zeta = 0. \tag{5}$$

In the above equations $\mathsf{P}(\boldsymbol{r}, t)$ is the pressure tensor, $\boldsymbol{q}(\boldsymbol{r}, t)$ is the heat flux, and $\zeta(\boldsymbol{r}, t)$ is the cooling rate due to energy dissipation in collisions. They are given by functionals of the one particle distribution function, so that the equations are purely formal until solving the Boltzmann equation for a particular situation. Let us point out that similar balance equations are obtained from the (exact) pseudo-Liouville equation which is valid for arbitrary density of the system [7]. The cooling rate ζ does not vanish for uniform systems, since dissipation in collisions is always present. As a consequence, simple inspection of (5) indicates that any state of a granular fluid must have gradients, be time dependent, or both. In particular, there are not uniform steady states.

It is important to stress the meaning of the temperature parameter T in granular fluids [3]. Due to the dissipative dynamics, the volume in phase space

of a freely evolving granular flow is not conserved, but is contracting. Therefore, the system can not reach an equilibrium state and there is no basis to formulate a thermodynamic theory for granular gases. For this reason, the expression "granular temperature" is often used, and in its definition as proportional to the velocity fluctuations, the Boltzmann constant is omitted.

The simplest state for a granular gas is the so-called *homogeneous cooling state* (HCS). In the case of a dilute gas, it is easily seen that the Boltzmann equation admits a solution of the form

$$f_H(\boldsymbol{v},t) = n_H v_H^{-d}(t)\phi\left(\frac{v}{v_H(t)}\right), \tag{6}$$

where

$$v_H(t) = \left[\frac{T_H(t)}{m}\right]^{1/2}. \tag{7}$$

The hydrodynamic fields associated to f_H verify

$$\boldsymbol{u}_H = 0, \quad \boldsymbol{\nabla} n_H = \boldsymbol{\nabla} T_H = 0, \quad T_H(t) = \frac{T_H(0)}{\left[1+\frac{\zeta_H(0)t}{2}\right]^2}. \tag{8}$$

This HCS plays the same role for granular flows as the equilibrium state for molecular fluids.

The standard procedure to derive explicit expressions for the fluxes in the balance equations (3)–(5) is the Chapman–Enskog procedure [15]. The method provides a normal solution of the Boltzmann equation as an expansion in the spatial gradients of the hydrodynamic fields. The results up to Navier–Stokes order have the form [8,16]

$$\mathsf{P}(\boldsymbol{r},t) = p\mathsf{I} - \eta\left[\boldsymbol{\nabla}\boldsymbol{u} + (\boldsymbol{\nabla}\boldsymbol{u})^{+} - \frac{2}{d}\mathsf{I}\boldsymbol{\nabla}\cdot\boldsymbol{u}\right], \tag{9}$$

$$\boldsymbol{q} = -\kappa\boldsymbol{\nabla} - \mu\boldsymbol{\nabla} n, \tag{10}$$

where p is the hydrostatic pressure, I the unit tensor, η is the shear viscosity, κ the thermal conductivity, and μ a new coefficient that has no analogue for elastic scattering. These transport coefficients depend on the coefficient of restitution α,

$$\eta = \eta^*(\alpha)\eta_0(T), \quad \kappa = \kappa^*(\alpha)\kappa_0(T), \quad \mu = \mu^*(\alpha)\frac{\kappa_0(T)T}{n}. \tag{11}$$

The quantities η_0 and κ_0 are the elastic limit values of the shear viscosity and thermal conductivity, respectively,

$$\eta_0 = \frac{2+d}{8}\Gamma(d/2)\pi^{-\frac{d-1}{2}}(mT)^{1/2}\sigma^{-(d-1)}, \tag{12}$$

$$\kappa_0 = \frac{d(d+2)^2}{16(d-1)}\Gamma(d/2)\pi^{-\frac{d-1}{2}}\left(\frac{T}{m}\right)^{1/2}\sigma^{-(d-1)}. \tag{13}$$

The dimensionless functions $\eta^*(\alpha)$, $\kappa^*(\alpha)$, and $\mu^*(\alpha)$ in the so-called first Sonine approximation read [8,16]

$$\eta^*(\alpha) = \left[\nu_1^*(\alpha) - \frac{\zeta^*(\alpha)}{2}\right]^{-1}, \tag{14}$$

$$\kappa^*(\alpha) = \left[\nu_2^*(\alpha) - \frac{2d}{d-1}\zeta^*(\alpha)\right]^{-1}[1 + c^*(\alpha)], \tag{15}$$

$$\mu^*(\alpha) = 2\zeta^*(\alpha)\left[\kappa^*(\alpha) + \frac{(d-1)c^*(\alpha)}{2d\zeta^*(\alpha)}\right]\left[\frac{2(d-1)}{d}\nu_2^*(\alpha) - 3\zeta^*(\alpha)\right]^{-1}, \tag{16}$$

where we have introduced the functions

$$\zeta^*(\alpha) = \frac{2+d}{4d}(1-\alpha^2)\left[1 + \frac{3}{32}c^*(\alpha)\right], \tag{17}$$

$$\nu_1^*(\alpha) = \frac{(3-3\alpha+2d)(1+\alpha)}{4d}\left[1 - \frac{1}{64}c^*(\alpha)\right], \tag{18}$$

$$\nu_2^*(\alpha) = \frac{1+\alpha}{d-1}\left[\frac{d-1}{2} + \frac{3(d+8)(1-\alpha)}{16} + \frac{4+5d-3(4-d)\alpha}{1024}c^*(\alpha)\right], \tag{19}$$

$$c^*(\alpha) = \frac{32(1-\alpha)(1-2\alpha^2)}{9+24d+(8d-41)\alpha+30\alpha^2(1-\alpha)}. \tag{20}$$

For consistency, the cooling rate ζ is needed up to second order in the gradients, since the hydrodynamic equations (3)–(5) depend on the gradients of the pressure tensor and the heat flux, which also are of second order. It is

$$\zeta(\boldsymbol{r},t) = \zeta^{(0)}(\boldsymbol{r},t) + \zeta^{(2)}(\boldsymbol{r},t). \tag{21}$$

The zeroth order in the gradients contribution is given by

$$\zeta^{(0)} = \zeta^*(\alpha)\frac{p}{\eta_0}, \tag{22}$$

¿From symmetry considerations it follows that $\zeta^{(2)}$ has the form

$$\begin{aligned}\zeta^{(2)} = {} & \zeta_1\nabla^2 T + \zeta_2\nabla^2 n + \zeta_3(\boldsymbol{\nabla}T)^2 + \zeta_4(\boldsymbol{\nabla}n)^2 + \zeta_5\boldsymbol{\nabla}n\cdot\boldsymbol{\nabla}T \\ & + \zeta_6(\boldsymbol{\nabla}\boldsymbol{u}):(\boldsymbol{\nabla}\boldsymbol{u}) + \zeta_7(\boldsymbol{\nabla}\boldsymbol{u}):(\boldsymbol{\nabla}\boldsymbol{u})^+.\end{aligned} \tag{23}$$

As long as I know, only the coefficients ζ_1 and ζ_2 of the linear terms have been calculated up to now [8,17]. For not very inelastic particles ($\alpha \geq 0.5$), their contributions to the transport equations are negligible as compared with those associated to κ and μ. A similar behaviour is expected for coefficients of the nonlinear terms in (23) so that the cooling rate can be accurately approximated by $\zeta^{(0)}$.

The derivation of the hydrodynamic equations sketched above is based on the assumption that there is a "normal solution" to the Boltzmann equation.

This solution is expected to describe the evolution of the system after an initial transient time in which the purely kinetic contributions decay. As mentioned in the Introduction, whether this "aging to hydrodynamics" actually occurs beyond the weakly inelasticity limit is not a trivial question [18,19]. Nevertheless, the validity of the hydrodynamic description in the linear region about the HCS has been verified for $\alpha \geq 0.65$ by comparing the numerical solutions of the Boltzmann equation and the hydrodynamic predictions [20]. Another example of applicability of the hydrodynamic equations, both in the linear and nonlinear regions, will be discussed below.

Granular fluids present quite peculiar steady states as compared with molecular fluids. Let us consider, for instance, the uniform shear flow state. It is characterized by a linear profile of the velocity flow and uniform density and temperature. In a molecular fluid, the temperature of the system increases monotonically in time due to viscous heating. On the other hand, (8) shows that in a granular flow the temperature changes in time due to the competition between the viscous heating (the term proportional to the square of the velocity gradient) and the cooling associated to energy dissipation in collisions. Then, a steady state is possible when both effects exactly compensate each other. This state has been extensively studied [5,21]. Another example of a granular steady state is a system confined between two boundaries at the same constant temperature or, similarly, between two walls that are vibrating in the same way. In the steady state the system is highly inhomogeneous [22] as compared with the equilibrium homogeneous state reached by a molecular fluid (away from the boundary layers). In the above two steady states, dissipation in collisions is coupled to the spatial gradients in the system, so that some dimensionless gradient is an increasing function on the coefficient of restitution. As a consequence, the limit of small gradients is equivalent to small inelasticity, and rheological effects play a very important role in granular fluids.

3 Linear Stability Analysis

One of the outstanding effects in granular fluids is the spontaneous development of vorticity fields and large scale inhomogeneities in density. The latter lead to the so-called clustering in which regions of high density coexist with almost empty regions. Clustering in freely evolving systems has been studied both theoretically and numerically [10,11], and it has been associated with an hydrodynamic instability against spatial perturbations of long wave length. Due to the monotonic cooling, freely evolving systems are very difficult to examine experimentally and clustering has not been observed in real systems under these conditions up to now. Nevertheless, very recently a series of experiments have been carried out in driven systems, and clustering has been clearly identified in them [23,24,25].

In order to investigate the nature of the clustering instability, let us start by considering a linear analysis about the HCS of the hydrodynamic equations. The linearization leads to partial differential equations with coefficients that are

independent of the position but depend on time. This follows from the time dependence of the reference state, the HCS. The time dependence of the coefficients can be eliminated by means of a change in the time and space variables, and a scaling of the hydrodynamic fields with those of the reference state. Then we define dimensionless space and time variables by

$$\tau = \frac{1}{2}\int_0^t dt' \nu_H(t'), \quad \boldsymbol{l} = \nu_H(t) v_H(t)^{-1} \frac{\boldsymbol{r}}{2}, \tag{24}$$

where $\nu_H = n_H T_H/\eta_0$ is a characteristic collision frequency for the HCS at time t. Note that the transformation of the space coordinate is time-independent. ¿From the evolution equation for the temperature (5) it is seen that in the τ scale

$$T_H(\tau) = T_H(0) e^{-2\zeta^* \tau}, \tag{25}$$

and using now the expression of T_H in (8) we get the relationship between the two time scales

$$e^{\tau} = \left(1 + \frac{t}{t_0}\right)^{1/\zeta^*}. \tag{26}$$

Denoting by

$$\delta y_i(\boldsymbol{l}, \tau) \equiv y_i(\boldsymbol{l}, \tau) - y_H(\tau) \tag{27}$$

the deviation of the hydrodynamic fields from their value in the HCS, the scaled fields are defined by

$$\theta = \frac{\delta T(\boldsymbol{l}, \tau)}{T_H(\tau)}, \quad \rho = \frac{\delta n(\boldsymbol{l}, \tau)}{n_H}, \quad \boldsymbol{\omega} = \frac{\delta \boldsymbol{u}(\boldsymbol{l}, \tau)}{v_H(\tau)}. \tag{28}$$

Next we introduce the Fourier transformed

$$\delta y_{\boldsymbol{k}i}(\tau) = \int d\boldsymbol{l}\, \exp(-i\boldsymbol{k}\cdot\boldsymbol{l}) \delta y_i(\boldsymbol{l}, \tau). \tag{29}$$

Now the equations for the Fourier hydrodynamic modes become ordinary differential equations with constant coefficients [8]

$$\frac{\partial}{\partial \tau}\rho_{\boldsymbol{k}} + ik\omega_{\boldsymbol{k}\|} = 0, \tag{30}$$

$$\left(\frac{\partial}{\partial \tau} - \zeta^* + \frac{d-1}{d}\eta^* k^2\right)\omega_{\boldsymbol{k}\|} + ik\theta_{\mathbf{k}} + ik\rho_{\boldsymbol{k}} = 0, \tag{31}$$

$$\left(\frac{\partial}{\partial \tau} - \zeta^* + \frac{1}{2}\eta^* k^2\right)\boldsymbol{\omega}_{\boldsymbol{k}\perp} = 0, \tag{32}$$

$$\left[\frac{\partial}{\partial \tau} + \zeta^* + \frac{d+2}{2(d-1)}\kappa^* k^2\right]\theta_{\boldsymbol{k}} + \left[2\zeta^* + \frac{d+2}{2(d-1)}\mu^* k^2\right]\rho_{\boldsymbol{k}} + \frac{2}{d}ik\omega_{\boldsymbol{k}\|} = 0. \tag{33}$$

The symbols $\omega_{\boldsymbol{k}\|}$ and $\omega_{\boldsymbol{k}\perp}$ denote the longitudinal and transversal components of the velocity field relative to the wave vector $\boldsymbol{k}$. Of course, for $d = 2$ the

transversal velocity $\omega_{\boldsymbol{k}\perp}$ only has one component. Equation (32) is decoupled from the others yielding

$$\boldsymbol{\omega}_{\boldsymbol{k}\perp}(\tau) = \boldsymbol{\omega}_{\boldsymbol{k}\perp}(0) e^{s_\perp \tau}, \tag{34}$$

where

$$s_\perp = \zeta^* - \frac{\eta^* k^2}{2}. \tag{35}$$

This identifies $d-1$ shear modes whose time evolution is governed by the shear viscosity. The time evolution of the other fields can be written in terms of the three eigenmodes whose frequencies $s_j(\alpha, k)$ are the solutions of the cubic algebraic equation

$$\begin{aligned} s^3 &+ \left[\frac{d+2}{2(d-1)}\kappa^* + \frac{d-1}{d}\eta^*\right] k^2 s^2 \\ &+ \left\{\frac{d+2}{2d}k^4\eta^*\kappa^* + \frac{k^2}{d}\left[d+2+(d-1)\zeta^*\eta^* - \frac{d(d+2)}{2(d-1)}\zeta^*\kappa^*\right] - \zeta^{*2}\right\} s \\ &+ \left[\frac{d+2}{2(d-1)}(\kappa^* - \mu^*)k^2 - \zeta^*\right] k^2 = 0. \end{aligned} \tag{36}$$

As an illustration, in Fig. 1 the dispersion relations $s_j = s_j(k)$, i.e. the solutions of (36), for a two-dimensional system with $\alpha = 0.9$ are shown. The results are similar for other values of α. It is seen that in the limit of very small k all the eigenvalues are real, so that there are not propagating or sound modes. For larger k, two of the modes become complex conjugates and only their (common) real part is plotted. Let us remind the meaning of the hydrodynamic modes. Given a small initial perturbation of the hydrodynamic fields, as long as the deviations of the fields from their values in the HCS stay small, the time evolution of the transversal component of the velocity is given by (34) while the other scaled fields obey expressions of the form

$$\delta y_{\boldsymbol{k}j}(\tau) = \sum_{i=1,2,\parallel} c_{ji}\varphi_i(\alpha, \boldsymbol{k}) e^{s_i(\alpha,k)\tau}, \tag{37}$$

where the functions $\varphi_i(\alpha, \boldsymbol{k})$ are given by the components of the eigenvectors of the matrix associated to the set of equations (30), (31), and (33) [26]. Three different regions are clearly identified in Fig. 1. They are delimited by the values $k_\perp$ and $k_\parallel$ of the wave number, where

$$k_\perp = \left(\frac{2\zeta^*}{\eta^*}\right)^{1/2}, \quad k_\parallel = \left[\frac{2(d-1)}{(d+2)(\kappa^* - \mu^*)}\right]^{1/2}. \tag{38}$$

For wave numbers in the range $k > k_\perp$, all the eigenvalues s_i have negative real parts. Therefore, perturbations of the scaled fields with values of k in this short wave length region always decay in time. The definitions (28) imply that also the perturbations of the actual hydrodynamic fields decay since the temperature of the HCS is a monotonic decreasing function of time.

In the wave number region $k_{\|} < k < k_{\perp}$, the eigenvalue $s_{\perp}$ is positive, indicating that the transversal component of the perturbation of the scaled velocity grows exponentially with τ. Therefore, vortices of the scaled velocity are expected to develop in time. A direct consequence of this result is that, after a time interval, the linearized hydrodynamic equations are not accurate to describe the time evolution of the system, even if the initial perturbation was small. Then, the complete nonlinear equations must be used. On the other hand, one can not conclude from here that the transversal component of the *actual* velocity field is linearly unstable. In fact, from (25) and (28) it is obtained

$$\delta \boldsymbol{u}_{\boldsymbol{k}\perp}(\tau) = \delta \boldsymbol{u}_{\boldsymbol{k}\perp}(0) \exp\left(-\frac{\eta^* k^2 \tau}{2}\right), \tag{39}$$

that decays monotonically in time.

In the long wave length region $k < k_{\|}$, the eigenvalue denoted by $s_{\|}$ is also positive, leading to perturbations of the scaled fields that grow in time. This mode contributes, in principle, to the density, the longitudinal component of the velocity field, and the temperature. In order to know the exact way in which it couples to each of these fields, we would have to determine the eigenfunctions of the linear problem, that are only known in the limit of α very close to unity. In any case, the same kind of analysis as presented above for the transversal component of the velocity, shows that $\delta u_{\boldsymbol{k}\|}$ and $\delta T_{\boldsymbol{k}}$ decay in time in the linear approximation, while $\delta n_{\boldsymbol{k}}$ grows in time, because it is scaled with the density of

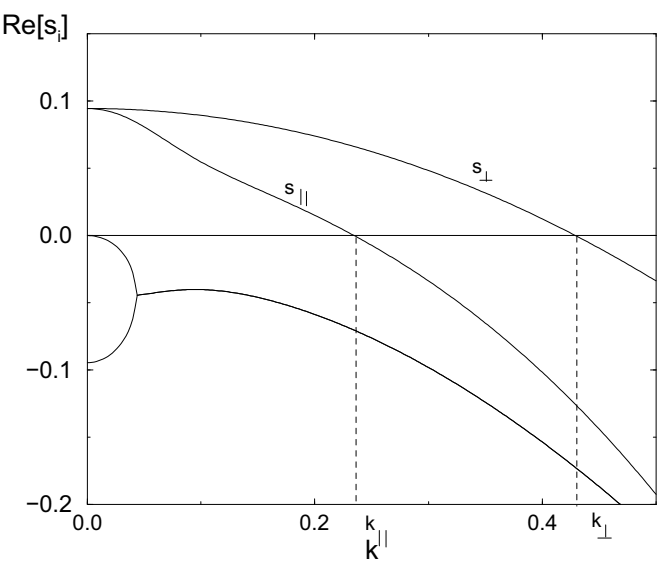

Fig. 1. Dispersion relations for the hydrodynamic modes vs the dimensionless wave number k for a two-dimensional granular gas. The frequencies s_j are measured in the dimensionless units defined in the text. The coefficient of restitution is $\alpha = 0.9$

the homogeneous state that does not depend on time. Nevertheless, even in this case it can not be concluded that the density is asymptotically unstable, since in the long time limit the linearized equations can not be applied.

4 The Shear Instability

In order to verify the validity of the hydrodynamic theory to describe the instability of a freely evolving granular fluid, we have used the direct simulation Monte Carlo (DSMC) method [27] to generate numerical solutions to the Boltzmann equation for a freely evolving granular gas. The DSMC is a very accurate and efficient tool to obtain numerical solutions of the Boltzmann equation. Since the details of the method for elastic systems are given in detail in [27] and the extension to inelastic collisions is trivial [28], it will not be discussed here.

The DSMC method is based on the Boltzmann equation, whose validity and accuracy to describe the problem at hand is not questioned in this paper. On the other hand, the method does not build any hydrodynamic description into the simulated system. Thus the comparison of theoretical predictions following from the hydrodynamic equations derived from the Boltzmann equation, and the numerical solutions obtained by means of the DSMC method is a test about the aging to hydrodynamics in the solutions of the inelastic Boltzmann equation. If, as it is the case in the situations we will discuss, the results also agree with Molecular Dynamics simulations, at least at a qualitative level, the whole picture provides a strong support to the validity of the complete theoretical scheme.

We have considered two-dimensional systems of size $L \times L$. The initial conditions for the simulations were an homogeneous distribution of particles and a Maxwellian distribution of velocities with a given temperature. Periodic boundary conditions were used in both directions. In order to characterize deviations from uniformity of mass and momentum distribution, the spatial average of the fluctuations of the number of particles density, $\Delta^* N$, and the ratio of the macroscopic kinetic energy to the thermal kinetic energy, E_K/E_T, were measured. These quantities are defined by

$$(\Delta^* N)^2 = \frac{1}{L^2} \int d\boldsymbol{r} \, \left[N(\boldsymbol{r}, t) - n_H\right]^2 , \tag{40}$$

$$E_K = \frac{1}{L^2} \int d\boldsymbol{r} \, \frac{m U^2(\boldsymbol{r}, t)}{2}, \quad E_T = \frac{1}{L^2} \int d\boldsymbol{r} \, E_T(\boldsymbol{r}, t), \tag{41}$$

where $N(\boldsymbol{r}, t)$, $\boldsymbol{U}(\boldsymbol{r}, t)$, and $E_T(\boldsymbol{r}, t)$ are the instantaneous number density, velocity, and internal energy fields. Of course, in the simulations the integrals are computed by some kind of coarse-graining.

Because of the periodic boundary conditions, the smallest allowed wave number is $k_m = 2\pi\ell/L$, where $\ell = 2v_H/\nu_H$ is the unit length used to define the dimensionless scale in (24). This quantity is proportional to the mean free path λ. If $k_m > k_\perp$ the system will be linearly stable, since all the linear perturbations

compatible with the boundary conditions will decay in time. We can estimate a critical size L_c for the onset of the shearing instability by

$$\frac{L_c}{\ell} = \frac{2\pi}{k_m} = \left(\frac{2\pi^2\eta^*}{\zeta^*}\right)^{1/2}. \tag{42}$$

This equations predicts a direct relationship between the restitution coefficient α and the critical size of the system measured in units of mean free path.

Given a value of α, we have run a series of simulations for increasing sizes of the system. The smallest size was chosen such that the system remained homogeneous along the simulation. The criteria for homogeneity were based on the time evolution of $\Delta^* N$ and E_K/E_T. For large systems, these quantities increase very quickly well above their initial homogeneous values, as consequence of spontaneous fluctuations in the system [29]. Let us point out that density structures, as measured by $\Delta^* N$ always show up once the separation between kinetic and thermal energy is quite large.

A visualization of the macroscopic velocity field for a system of size $L = 20\lambda$ is shown in Fig. 2. The coefficient of restitution is $\alpha = 0.8$. Time measured in the reduced units defined in Eq. (24) is $\tau = 96.3$. Fully developed velocity vortices are clearly observed, indicating that the size of the system is larger than the critical value L_c. A density contour plot for the same system at the same time is presented in Fig. (3). The lighter areas correspond to higher densities and the darker ones to lower densities. The density interval is $0.75 < n/n_H < 1.25$, i.e. no significant clustering is present yet, although the simulations showed that density inhomogeneities are growing very fast.

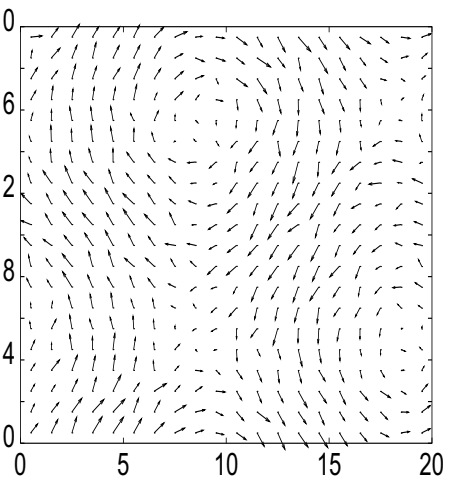

Fig. 2. Instantaneous macroscopic velocity field in a two-dimensional freely evolving system. Time measured from the initial homogeneous configuration in reduced units is $\tau = 96.3$. Length is measured in units of mean free path λ. The coefficient of restitution is $\alpha = 0.8$

A comparison of the hydrodynamic fields in Figs. 2 and 3 indicates that the velocity field and the density fields have similar structures. High density regions are developing at the boundary of the velocity vortices, while low density regions are located at their centers. This is another indication of the close connection between velocity and density instabilities even in the initial set up.

By decreasing L until reaching a value for which the system stay homogeneous for the duration of the simulation, the critical value $L_c(\alpha)$ can be bounded. In Fig. 4 the results for several values of the coefficient of restitution in the interval $0.7 \leq \alpha \leq 0.95$ are plotted (circles). The solid line is the theoretical prediction from the hydrodynamic stability analysis, given by (42). In the elastic limit, $\alpha \to 1$, the critical length diverges since the system is stable even in the thermodynamic limit. The agreement is fairly good over the entire range of inelasticity considered.

5 Clustering Instability

Although the physical origin of the density clustering can be easily understood by qualitative arguments as indicated in the Introduction, translating it into a more detailed and quantitative theory is not a trivial task. In fact, two different mechanisms have been proposed in the literature as responsible for the initial density nonuniformity. The most direct explanation is that it is produced as a linear hydrodynamic effect associated with a fluctuation of the mode governing the linear evolution of the longitudinal component of the flow field $u_{\mathbf{k}\|}$ [30,31]. Another possibility starts with a fluctuation of the transversal component of the velocity, i.e. the shear mode. Because of the growth of the fluctuation of the scaled transversal velocity, nonlinear effects become important after some transient period of time. In particular, the nonlinear coupling of the temperature with the shear mode leads to an increase of the temperature in the regions of

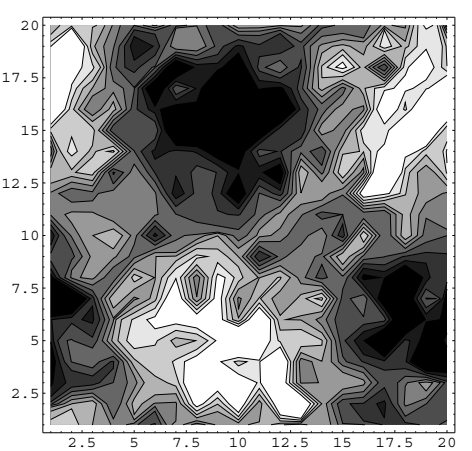

Fig. 3. Density contour plot for the same system as in Fig. 2. Lighter areas correspond to higher densities

larger vorticity. Then, a pressure gradient shows up and produces the density initial density fluctuation from which the cluster develops.

The two clustering mechanisms discussed above are physically different. While in the linear scenario velocity vortices and density clustering appear as independent non-correlated effects, in the nonlinear one they are closely related. The numerical results discussed in the previous Section support this latter possibility. Firstly, we saw that the velocity vortices always show up before the density clusters. Secondly, and more important, the structure of the initial formation of density inhomogeneities is the same as that of the velocity vortices, with the higher density regions located in the borders of the vortices.

In the following we will present DSMC results obtained in a three-dimensional system. The dispersion relations are similar to those for a two-dimensional system illustrated in Fig. 1. The spheres were confined between two infinite plates perpendicular to the x axis and separated a distance L. Periodic boundary conditions were applied at the plates, while the system was simulated homogeneous in the other two directions. This was accomplished by assuming uniform distribution along the y and z axis at all times.

Starting from an homogeneous freely evolving system, i.e. a granular gas in the HCS, an instantaneous external perturbation of the flow velocity given by

$$\boldsymbol{\omega}(x,0) = \boldsymbol{\omega}_0 \sin(q_0 x) \tag{43}$$

was introduced. The wave number is $q_0 = 2\pi/L$ and the amplitude $\boldsymbol{\omega}$ has components $\omega_{0x} = \omega_{0y} = \omega_{0z} \equiv \omega_0$. For this velocity perturbation ω_y and ω_z are the transversal components, while ω_x is the longitudinal one. This is because only gradients in the x direction are allowed.

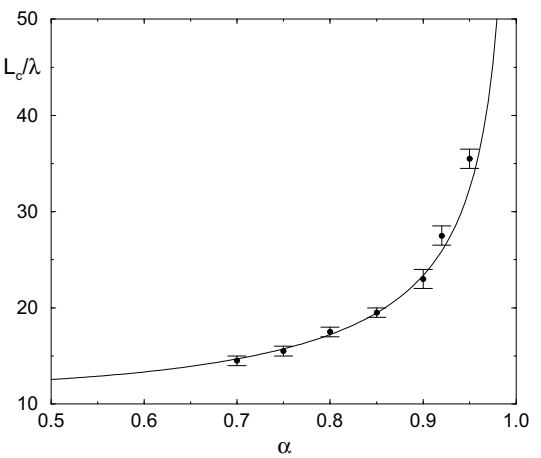

Fig. 4. Critical size L_c, measured in units of mean free path λ, for the shearing instability in a system of hard disks as a function of the coefficient of restitution α

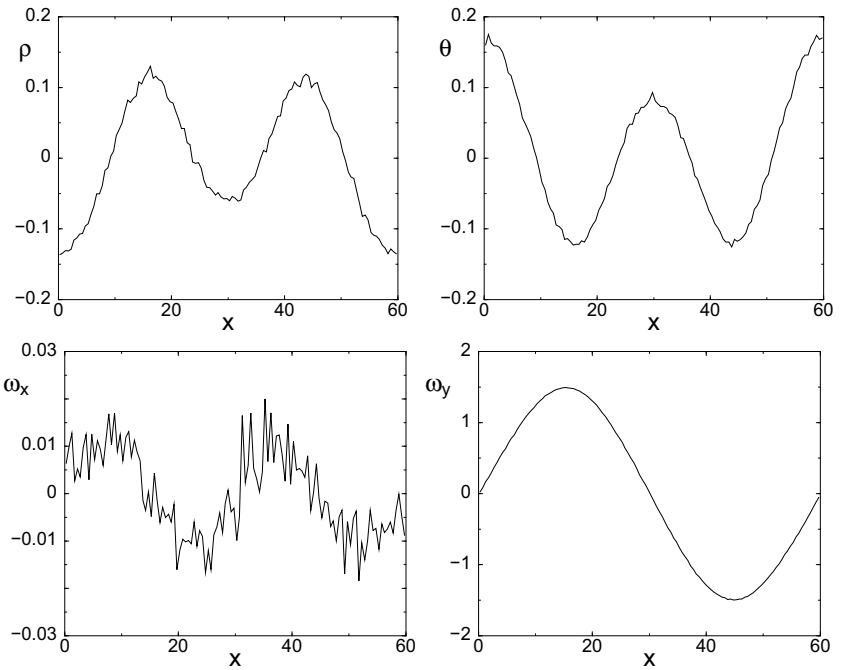

Fig. 5. Hydrodynamic profiles at a reduced time $\tau = 148.25$ following an harmonic perturbation of the three components of velocity field in a three-dimensional granular gas. The field are scaled as defined in the text. The distance x is measured in units of mean free path λ

An example of the shape of the hydrodynamic profiles after a velocity perturbation (43) with $\omega_0 = 0.05\sqrt{2}$ was introduced is given in Fig. 5. The coefficient of restitution is $\alpha = 0.95$, and the time in reduced units is $\tau = 148.5$. The size of the system is $L = 60\lambda$, where λ is the mean free path. For the above values of L and α the maximum allowed wave number in reduced units is $k_m \simeq 0.164$ and from the dispersion relations it is seen that $k_m < k_\parallel < k_\perp$ and $k_\perp < 2k_m$ [32]. Therefore, while the longitudinal and transversal modes of the scaled fields corresponding to Fourier components with wave number k_m are linearly unstable, those corresponding to jk_m, $j = 2, 3, 4, \cdots$, decay in the linear approximation and their eventual growth can be due only to nonlinear hydrodynamic couplings.

The profiles in Fig. 5 indicate that, for the time considered, ω_y is still well described by the lowest harmonic, i.e. by the linear approximation. In fact, the analysis of the time evolution of the amplitude of the perturbation shows that it is very well fitted by the linear law (34). Note that the amplitude of the perturbation of ω_y has already increased by a factor of the order of twenty. The other transversal component of the velocity ω_z shows a similar behavior and, therefore, has not been plotted. On the other hand, the density, the temperature, and the longitudinal component of the velocity exhibit quite different profiles. A simple glance clearly indicates that they are dominated by the second harmonic.

Moreover, by following their time evolution it is seen that the amplitudes are increasing in time. As discussed above, both features can not be explained by the linear theory and indicate the dominant role played by nonlinear effects. Besides, the difference in magnitude between the transversal velocity field and the other hydrodynamic fields, strongly indicates that the transversal modes have enslaved the other ones through some nonlinear coupling. This property has been used to present a simple mode coupling theory that leads to a good agreement with the DSMC results. A detailed account of the theory and the comparison with the simulations results is given in [32].

The analysis of the clustering instability presented here is limited to a low density gas described by the Boltzmann equation. Therefore, once the system develops very strong density inhomogeneities our description is no longer valid. That means that our conclusions are restricted to the initial setup of the clustering. For latter times, a more general theory incorporating excluded volume effects is needed. In this context, an interesting open question is whether the density in the cluster is limited by purely geometrical reasons or there is a lower saturation density, as seems to indicate the DSMC results and also some Molecular Dynamics simulations [33].

Acknowledgments

The material presented here is based in work carried out in collaboration with J.W. Dufty, M.J. Ruiz–Montero, F. Moreno, and D. Cubero to whom I am greatly indebted. The research was partially supported by the Dirección General de Investigación Científica y Técnica (Spain) through Grant No. PB98–1124.

References

1. H.M. Jaeger, S.R. Nagel, R.P. Behringer: Rev. Mod. Phys. **68**, 1259 (1996)
2. P.K. Haff: J. Fluid Mech. **83**, 401 (1983)
3. C.S. Campbell: Annu. Rev. Fluid Mech. **22**, 57 (1990)
4. J.T. Jenkins, S. Savage: J. Fluid Mech. **130**, 187 (1983); J.T. Jenkins, M.W. Richmann: Arch. Rat. Mech. **87**, 355 (1985)
5. C. Lun, S. Savage, D. Jeffrey, N. Chepurniy: J. Fluid Mech. **140**, 223 (1984)
6. For a recent review and references see, for instance, M. Ernst: 'Kinetic Theory of Granular Fluids'. In *Dynamics: Models and Kinetic Methods for Non-equilibrium Many Body Systems*, ed. by J. Karkheck (NATO ASI Series, Kluver, Dordrecht 2000)
7. J.J. Brey, J.W. Dufty, A. Santos: J. Stat. Phys. **87**, 1051 (1997); T.P.C. van Noije, M.H. Ernst: Physica A **251**, 266 (1998)
8. J.J. Brey, J.W. Dufty, C.S. Kim, A. Santos: Phys. Rev. E **58**, 4638 (1998)
9. N. Sela, I. Goldhirsch: J. Fluid Mech. **361**, 41 (1998)
10. I. Goldhirsch, G. Zanetti: Phys. Rev. Lett. **70**, 1619 (1993); I. Goldhirsch, M.L. Tan, G. Zanetti: J. Sci. Comput. **8**, 1 (1993)
11. S. McNamara, W.R. Young: Phys. Rev. E **50**, R28 (1994); Phys. Rev. E **53**, 5089 (1996)

12. F.G. Bridges, A. Hatzes, D.N.C. Lin: Nature **309**, 333 (1984); R.M. Brach: J. Appl. Mech. **56**, 133 (1989)
13. G. Kuwabara and K. Kono: Jpn. J. Appl. Phys. Part 1, **26**, 1230 (1987); N. Brilliantov, F. Spahn, J.M. Hertzsch, T. Pöschel: Phys. Rev. E **53** 5382 (1996); T. Schwager, T. Pöschel: Phys. Rev. E **57**, 650 (1998)
14. A. Goldsthein, M. Shapiro, J. Fluid Mech. **282**, 75 (1995)
15. J.A. McLennan: *Introduction to Nonequilibrium Statistical Mechanics* (Prentice–Hall, Englewood Cliffs, NJ 1989)
16. J.J. Brey and D. Cubero: 'Hydrodynamic transport coefficients in granular gases". In: *Granular Gases*, ed. by S. Luding, T. Pöschel (Springer-Verlag, Berlin 2000) pp. 59–78.
17. D. Cubero: *Ph. D. thesis* Universidad de Sevilla, 2000
18. L.P. Kadanoff: Rev. Mod. Phys. **71**, 435 (1999)
19. M.L. Tan, I. Goldhirsch: Phys. Rev. Lett. **81**, 3022 (1998)
20. J.J. Brey, M.J. Ruiz–Montero, D. Cubero: Europhys. Lett. **48**, 359 (1999)
21. J.T. Jenkins, M.W. Richman; J. Fluid Mech. **192**, 313 (1988) C.S. Campbell: J. Fluid Mech. **203**, 449 (1989) M.A. Hopkins, H.H. Shen: J. Fluid Mech. **244**, 477 (1992) N. Sela, I. Goldhirsch, S.H. Noskowicz: Phys. Fluids **8**, 2337 (1996) J.J. Brey, M.J. Ruiz–Montero, F. Moreno: Phys. Rev. E **55**, 2846 (1997) J.M. Montanero, V. Garzó, A. Santos, J.J. Brey: J. Fluid Mech. **99**, 391 (1999)
22. E.L. Grossman, T. Zhou, E. Ben–Naim: Phys. Rev. E **55**, 4200 (1997) J.J. Brey, D. Cubero: Phys. Rev. E, **98**, 57 (1998) J.J. Brey, D. Cubero, F. Moreno, M.J. Ruiz–Montero: unpublished
23. W. Losert, D.G.W. Cooper, J.P. Gollup: Phys. Rev. E **59**, 5855 (1999)
24. J.S. Olafsen, J.S. Urbach: Phys. Rev. Lett. **81**, 4369 (1998) Phys. Rev. E **60**, R2468 (1999)
25. R. Cafiero, S. Luding, H.J. Herrmann: Phys. Rev. Lett. **84**, 6014 (2000)
26. P. Résibois and M. de Leener: *Classical Kinetic Theory of Fluids*, (John Wiley, New York 1977)
27. G. Bird: *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, (Clarendon, Oxford 1994)
28. J.J. Brey, M.J. Ruiz–Montero, D. Cubero: Phys. Rev. E **54**, 3664 (1996)
29. J.J. Brey, M.J. Ruiz–Montero, F. Moreno: Phys. Fluids **10**, 2976 (1998)
30. P. Deltour, J.L. Barrat: J. Phys. I **7**, 137 (1997)
31. R. Brito, M.H. Ernst: Europhys. Lett. **43**, 497 (1998)
32. J.J. Brey, M.J. Ruiz–Montero, D. Cubero: Phys. Rev. E **60**, 3150 (1999)
33. See, for instance, Fig. 9 in the second paper mentioned in Ref. [11]

Granular Gases – The Early Stage

Nikolai V. Brilliantov[1,2] and Thorsten Pöschel[3]

[1] MPIKG, Am Mühlenberg, 14424, Potsdam, Germany
[2] Moscow State University, Physics Department, Moscow 119899, Russia
[3] Humboldt-University Berlin - Charité, Institute for Biochemistry,
Hessische Str. 3-4, D-10115 Berlin, Germany. e-mail: thorsten@physik.hu-berlin.de
http://summa.physik.hu-berlin.de/∼thorsten/

Abstract. We investigate the evolution of the velocity distribution function of a granular gas composed of viscoelastic particles in the homogeneous cooling state, i.e. before clustering occurs. The deviation of the velocity distribution function from the Maxwellian distribution is quantified by a Sonine polynomials expansion. The first non-vanishing Sonine coefficient $a_2(t)$, reveals a complex time dependence which allows to assign the granular gas the property of an age. We discuss the possibility to measure the age of a granular gas.

1 Introduction

Granular Gases as rarefied systems of granular particles in the absence of gravity may be exemplified by a cloud of interstellar dust. Similar as molecular gases Granular Gases may be described within the concepts of classical Statistical Mechanics, such as temperature T, velocity distribution function $f(v)$, etc. Once initialized with a certain velocity distribution, Granular Gases cool down due to inelastic collisions of their particles. Although these systems are extremely simple, in principle, they reveal a variety of structure formation and much work has been done recently to characterize the properties of cooling Granular Gases (see [1] with many references therein).

Most of these results have been obtained under the assumption that the coefficient of restitution ϵ which characterizes the loss of energy of two colliding particles (see below), is a material constant. The assumption $\epsilon = \text{const.}$, however, does not only contradict experiments which show that ϵ depends significantly on the impact velocity [2,3], but it contradicts even some basic mechanical laws [4]. The simplest physically correct description of dissipative particle collisions is based on the assumption of viscoelastic material deformation during collisions [5], which is valid for particle collisions in a certain range of impact velocity and is in good agreement with experimental data [3,6].

We investigate the statistical properties of Granular Gases of viscoelastic particles for which the dependence of the restitution coefficient on the impact velocity $\epsilon = \epsilon(v_{\text{imp}})$ is known [4,7]. Starting from a homogeneous distribution we study the early stage of its evolution, where no spatial structures, as clusters [8,9] and vortexes [10], have emerged yet. This stage is called the homogeneously cooling state. Our results show that the properties of Granular Gases change qualitatively if one takes into account viscoelastic material properties,

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 408–419, 2001.
© Springer-Verlag Berlin Heidelberg 2001

i.e. $\epsilon = \epsilon(v_{\rm imp})$ [11], as compared with the equivalent system, but under the oversimplified assumption $\epsilon = \text{const.}$.

In the next Section we discuss briefly the impact-velocity dependence of the normal restitution coefficient. In Sec. 3 we introduce the method to study the velocity distribution by means of the Boltzmann equation with the Sonine polynomials expansion formalism, and discuss the results obtained for gases of particles interacting with a constant restitution coefficient. Our main results, which describe the time evolution of the velocity distribution for granular gases of viscoelastic particles are derived in Sec. 4. In Sec. 5 we discuss the concept of the age of a granular gas which is based on the time evolution of the velocity distribution. Finally, in conclusion we summarize our findings.

2 Two-Particle Interaction of Viscoelastic Spheres

The microscopic dynamics of granular particles is governed by the (normal) restitution coefficient ϵ which relates the normal components of the particle velocities before and after a collision, $\boldsymbol{v}_{ij} \equiv \boldsymbol{v}_i - \boldsymbol{v}_j$ and $\boldsymbol{v}^*_{ij} \equiv \boldsymbol{v}^*_i - \boldsymbol{v}^*_j$ by $\left|\boldsymbol{v}^*_{ij}\boldsymbol{e}\right| = \epsilon \left|\boldsymbol{v}_{ij}\boldsymbol{e}\right|$. The unit vector $\boldsymbol{e} = \boldsymbol{r}_{ij}/\left|\boldsymbol{r}_{ij}\right|$ gives the direction of the inter-center vector $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$ at the instant of the collision. From the conservation of momentum one finds the change of velocity for the colliding particles:

$$\boldsymbol{v}^*_i = \boldsymbol{v}_i - \frac{1}{2}\left(1+\epsilon\right)\left(\boldsymbol{v}_{ij}\cdot\boldsymbol{e}\right)\boldsymbol{e}\,, \qquad \boldsymbol{v}^*_j = \boldsymbol{v}_j + \frac{1}{2}\left(1+\epsilon\right)\left(\boldsymbol{v}_{ij}\cdot\boldsymbol{e}\right)\boldsymbol{e}\,. \tag{1}$$

For elastic collisions one has $\epsilon = 1$ and for inelastic collisions ϵ decreases with increasing degree of inelasticity.

In literature it is frequently assumed that the restitution coefficient is a material constant, $\epsilon = \text{const}$. Experiments, e.g. [2,3], as well as theoretical investigations [5] show, however, that this assumption is not consistent with the nature of inelastic collisions, it does not agree even with a dimension analysis [4]. The impact velocity dependence of the restitution coefficient $\epsilon\left(v_{\rm imp}\right) = \epsilon\left(\left|\boldsymbol{e}\,\boldsymbol{v}_{ij}\right|\right)$ has been obtained by generalizing Hertz's contact problem to viscoelastic spheres [5]. From the generalized Hertz equation [12] one obtains the velocity-dependent restitution coefficient for viscoelastic spheres [7]

$$\epsilon = 1 - C_1 A \alpha^{2/5}\left|\boldsymbol{e}\,\boldsymbol{v}_{ij}\right|^{1/5} + C_2 A^2 \alpha^{4/5}\left|\boldsymbol{e}\,\boldsymbol{v}_{ij}\right|^{2/5} \mp \cdots \tag{2}$$

with

$$\alpha = \left(\frac{3}{2}\right)^{3/2} \frac{Y\sqrt{R^{\rm eff}}}{m^{\rm eff}\left(1-\nu^2\right)}\,, \tag{3}$$

where Y is the Young modulus, ν is the Poisson ratio, and A depends on dissipative parameters of the particle material (for details see [5]). The effective mass and radius are defined as

$$R^{\rm eff} = R_1 R_2/(R_1+R_2) \qquad\qquad m^{\rm eff} = m_1 m_2/(m_1+m_2) \tag{4}$$

with $R_{1/2}$ and $m_{1/2}$ being the radii and masses of the colliding particles. The constants are given by [4,7]

$$C_1 = \frac{\Gamma(3/5)\sqrt{\pi}}{2^{1/5}5^{2/5}\Gamma(21/10)} \approx 1.1534 \qquad C_2 = \frac{3}{5}C_1^2 \approx 0.7982\,. \tag{5}$$

Equation (2) refers to the case of pure viscoelastic interaction, i.e. when the relative velocity $|\boldsymbol{v}_{ij}\boldsymbol{e}|$ is not too large (to avoid plastic deformation of the particles) and is not too small (to allow to neglect surface effects such as roughness, adhesion and van der Waals interactions). The dependence of $\epsilon = \epsilon\,(|\boldsymbol{e}\boldsymbol{v}_{ij}|)$ (without the material dependence) was already mentioned in [6] where heuristic arguments have been applied.

In what follows we consider a granular gas of identical viscoelastic spheres of unit mass.

3 Kinetics of Granular Gases: The Case $\epsilon = \text{const.}$

The evolution of the velocity distribution function is generally described by the Boltzmann-Enskog equation, which for the force-free case reads [13,14,15]:

$$\begin{aligned}\frac{\partial}{\partial t}f\,(\boldsymbol{v},t) = g_2(\sigma)\sigma^2 \int d\boldsymbol{v}_2 \int d\boldsymbol{e}\Theta\,(-\boldsymbol{v}_{12}\cdot\boldsymbol{e})\,|\boldsymbol{v}_{12}\cdot\boldsymbol{e}| \\ \times\,\{\chi f(\boldsymbol{v}_1^{**},t)f(\boldsymbol{v}_2^{**},t) - f(\boldsymbol{v}_1,t)f(\boldsymbol{v}_2,t)\} \equiv g_2(\sigma)I(f,f)\,,\end{aligned} \tag{6}$$

where $\sigma = 2R$ is the diameter of particles, $g_2(\sigma) = (2-\eta)/2(1-\eta)^3$ ($\eta = \frac{1}{6}\,\pi n\sigma^3$ is packing fraction) denotes the contact value of the two-particle correlation function [16], which accounts for the increasing collision frequency due to the excluded volume effects; $\Theta(x)$ is the Heaviside step-function. The velocities $\boldsymbol{v}_1^{**}$ and $\boldsymbol{v}_2^{**}$ refer to the precollisional velocities of the so-called inverse collision, which results with $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ as the after-collisional velocities. Finally the factor χ in the gain term appears respectively from the Jacobian of the transformation $d\boldsymbol{v}_1^{**}d\boldsymbol{v}_2^{**} \to d\boldsymbol{v}_1 d\boldsymbol{v}_2$ and from the relation between the lengths of the collisional cylinders $\epsilon\,|\boldsymbol{v}_{12}^{**}\cdot\boldsymbol{e}|\,dt = |\boldsymbol{v}_{12}\cdot\boldsymbol{e}|\,dt$. [17]. For the constant restitution coefficient $\epsilon = \text{const.}$ this reads $\chi = \epsilon^{-2}$ [14,18], i.e., it is independent on the impact velocity and, therefore, independent on time.

With the scaling Ansatz for the distribution function

$$f(\boldsymbol{v},t) = \frac{n}{v_0^3(t)}\tilde{f}\left(\frac{v}{v_0(t)}\right) = \frac{n}{v_0^3(t)}\tilde{f}(\boldsymbol{c})\,, \tag{7}$$

where n is the number density of the granular gas and $v_0(t)$ is the thermal velocity defined by

$$\frac{3}{2}nT(t) = \int d\boldsymbol{v}\frac{v^2}{2}f(\boldsymbol{v},t)\,, = \frac{3}{2}nv_0^2(t) \tag{8}$$

the Boltzmann equation may be reduced to two independent equations: one for the (time-independent) scaling function $\tilde{f}(\boldsymbol{c})$, and the other one for the time-dependence of the thermal velocity (i.e. for the temperature). Solving the equation for the temperature one obtains $T(t) = T_0/\left(1 + t/\tau\right)^2$ [13].

The solution of the other equation for $\tilde{f}(\boldsymbol{c})$, may be found in terms of the Sonine polynomial expansion [14,18,19]. For the case of $\epsilon = \text{const.}$ it reads

$$\tilde{f}(\boldsymbol{c}) = \phi(c) \left\{ 1 + \sum_{p=1}^{\infty} a_p S_p(c^2) \right\} , \tag{9}$$

where $\phi(c) \equiv \pi^{-d/2} \exp(-c^2)$ is the Maxwellian distribution for the rescaled velocity, and the first few Sonine polynomials are

$$S_0(x) = 1, \quad S_1(x) = -x^2 + \frac{3}{2}, \quad S_2(x) = \frac{x^2}{2} - \frac{5x}{2} + \frac{15}{8} . \tag{10}$$

The leading (zero-order) term in Eq. (9) is the Maxwellian distribution, while the next-order terms, characterized by the coefficients a_i describe the deviation of the distribution from the Maxwellian. For the case of $\epsilon = \text{const.}$ the velocity distribution function $\tilde{f}(\boldsymbol{c})$ is time-independent, therefore, the coefficients of the Sonine polynomials expansion are constants.

If the inelasticity is small, one can restrict to the first non-vanishing term beyond the Maxwellian which has the coefficient a_2 ($a_1 = 0$, according to the definition of temperature [14,18,19]). For $\epsilon = \text{const.}$ this coefficient reads [20]

$$a_2 = \frac{16(1-\epsilon)(1-2\epsilon^2)}{9 + 24d + 8\epsilon d + 41\epsilon + 30(1-\epsilon)\epsilon^2} . \tag{11}$$

A more accurate expression for a_2 may be found in [19].

4 Kinetics of Granular Gases: Viscoelastic Particles

For viscoelastic particles the restitution coefficient ϵ depends on the impact velocity due to Eq. (2). Hence, the factor χ in the Boltzmann equation (6) is not anymore constant as for $\epsilon = \text{const.}$ but it reads

$$\chi = 1 + \frac{11}{5} C_1 A \alpha^{2/5} \left| \boldsymbol{v}_{12} \cdot \boldsymbol{e} \right|^{1/5} + \frac{66}{25} C_1^2 A^2 \alpha^{4/5} \left| \boldsymbol{v}_{12} \cdot \boldsymbol{e} \right|^{2/5} + \cdots \tag{12}$$

Using again the scaling Ansatz (7) the rhs of the Boltzmann equation does not factorize into two parts, one depending only on the scaling function $\tilde{f}(\boldsymbol{c})$, and another one, depending only on v_0. Consequently, one can not obtain a set of decoupled equations for $\tilde{f}(\boldsymbol{c})$ and for the temperature.

Nevertheless, it is worth to substitute the generalized scaling Ansatz

$$f(\boldsymbol{v}, t) = \frac{n}{v_0^3(t)} \tilde{f}(\boldsymbol{c}, t) \tag{13}$$

into the kinetic equation (6). After some algebra Eq.(6) may be recast into the form

$$\frac{\mu_2}{3}\left(3+c_1\frac{\partial}{\partial c_1}\right)\tilde{f}(\boldsymbol{c},t)+B^{-1}\frac{\partial}{\partial t}\tilde{f}(\boldsymbol{c},t)=\tilde{I}\left(\tilde{f},\tilde{f}\right) \tag{14}$$

where we define the dimensionless collisional integral

$$\tilde{I}\left(\tilde{f},\tilde{f}\right)=\int d\boldsymbol{c}_2\int d\boldsymbol{e}\,\Theta(-\boldsymbol{c}_{12}\cdot\boldsymbol{e})|\boldsymbol{c}_{12}\cdot\boldsymbol{e}|\left\{\tilde{\chi}\tilde{f}(\boldsymbol{c}_1^{**},t)\tilde{f}(\boldsymbol{c}_2^{**},t)-\tilde{f}(\boldsymbol{c}_1,t)\tilde{f}(\boldsymbol{c}_2,t)\right\} \tag{15}$$

with the reduced factor $\tilde{\chi}$

$$\tilde{\chi}=1+\frac{11}{5}C_1\delta'\left|\boldsymbol{c}_{12}\cdot\boldsymbol{e}\right|^{1/5}+\frac{66}{25}C_1^2\delta'^{\,2}\left|\boldsymbol{c}_{12}\cdot\boldsymbol{e}\right|^{2/5}+\cdots \tag{16}$$

which depends now on time via a quantity

$$\delta'(t)\equiv A\alpha^{2/5}\left[2T(t)\right]^{1/10}\equiv\delta\left[2T(t)/T_0\right]^{1/10} \tag{17}$$

Here $\delta\equiv A\alpha^{2/5}[T_0]^{1/10}$ and T_0 is the initial temperature. We also define $B=B(t)\equiv v_0(t)g_2(\sigma)\sigma^2 n$, and the moments of the dimensionless collision integral

$$\mu_p\equiv-\int d\boldsymbol{c}_1 c_1^p\tilde{I}\left(\tilde{f},\tilde{f}\right)\ . \tag{18}$$

According to the definitions (8), (18) and of B, the second moment μ_2 defines the evolution of the temperature:

$$\frac{dT}{dt}=\frac{1}{3}g_2(\sigma)\sigma^2 n v_0^3\int d\boldsymbol{c}_1 c_1^2\tilde{I}\left(\tilde{f},\tilde{f}\right)=-\frac{2}{3}BT\mu_2\,. \tag{19}$$

The velocity distribution function we again describe by a Sonine polynomials expansion as introduced in Eq. (9). Since in contrast to the case $\epsilon=$ const., the Boltzmann equation for a gas of viscoelastic particles does not factorize into a time dependent equation for temperature and a time-independent equation for the velocity distribution, the Sonine coefficients a_i are not constants but depend explicitely on time, i.e. one has now

$$\tilde{f}(\boldsymbol{c},t)=\phi(c)\left\{1+\sum_{p=1}^{\infty}a_p(t)S_p(c^2)\right\}\,. \tag{20}$$

Equations for $a_p(t)$ may be found by multiplying both sides of Eq.(14) with c_1^p and integrating over $d\boldsymbol{c}_1$. One obtains

$$\frac{\mu_2}{3}p\left\langle c^p\right\rangle-B^{-1}\sum_{k=1}^{\infty}\dot{a}_k\nu_{kp}=\mu_p \tag{21}$$

where integration by parts has been performed and where we define

$$\nu_{kp}\equiv\int\phi(c)c^p S_k(c^2)d\boldsymbol{c}\,;\quad\left\langle c^p\right\rangle\equiv\int c^p\tilde{f}(\boldsymbol{c},t)d\boldsymbol{c}\,. \tag{22}$$

From (21) we see already that the granular temperature and the Sonine coefficients and, hence, the distribution function do not evolve independently.

The calculation of ν_{kp} is straightforward; the first few of them read $\nu_{22} = 0$, $\nu_{24} = \frac{15}{4}$. The odd moments $\langle c^{2n+1} \rangle$ are zero, while the even ones, $\langle c^{2n} \rangle$ may be expressed in terms of a_k with $0 \leq k \leq n$. Calculations show that $\langle c^2 \rangle = \frac{3}{2}$, implying $a_1 = 0$, according to the definition of temperature (8) (e.g.[14]), and that $\langle c^4 \rangle = \frac{15}{4}(1 + a_2)$.

Now we assume that the dissipation is not large so that the deviation from the Maxwellian distribution may be sufficiently described by the second term in the expansion (9) only, with all higher-order terms with $p > 2$ discarded. Then (21) is an equation for the coefficient a_2. Using the above results for ν_{22}, ν_{24}, $\langle c^2 \rangle$ and $\langle c^4 \rangle$ it is easy to show that Eq.(21) converts for $p = 2$ into identity, while for $p = 4$ it reads:

$$\dot{a}_2 - \frac{4}{3} B\mu_2 (1 + a_2) + \frac{4}{15} B\mu_4 = 0 \,. \tag{23}$$

With the approximation $\tilde{f} = \phi(c)[1+a_2(t)S_2(c^2)]$ the time-dependent coefficients $\mu_p(t)$ may be expressed in terms of a_2 due to the definition (18). Using the properties of the collision integral (e.g. the conservation of the total momentum at collision) one can obtain relations for the $\mu_p(t)$ (e.g. [14]):

$$\mu_p = -\frac{1}{2} \int d\boldsymbol{c}_1 \int d\boldsymbol{c}_2 \int d\boldsymbol{e} \Theta(-\boldsymbol{c}_{12} \cdot \boldsymbol{e}) \, |\boldsymbol{c}_{12} \cdot \boldsymbol{e}| \, \phi(c_1)\phi(c_2) \times \\ \left\{1 + a_2 \left[S_2(c_1^2) + S_2(c_2^2)\right] + a_2^2 \, S_2(c_1^2) S_2(c_2^2)\right\} \Delta(c_1^p + c_2^p)$$

where $\Delta\psi(\boldsymbol{c}_i) \equiv [\psi(\boldsymbol{c}_i^*) - \psi(\boldsymbol{c}_i)]$ denotes the change of some function $\psi(\boldsymbol{c}_i)$ in a direct collision. Calculations up to the second order in δ yield [17]

$$\mu_2 = \sum_{k=0}^{2} \sum_{n=0}^{2} \mathcal{A}_{kn} \delta'^{\,k} a_2^n \tag{24}$$

where the coefficients $\mathcal{A}$ are pure numbers: $\mathcal{A}_{00} = 0$, $\mathcal{A}_{01} = 0$, $\mathcal{A}_{02} = 0$, $\mathcal{A}_{10} = \omega_0$, $\mathcal{A}_{11} = \frac{6}{25}\omega_0$, $\mathcal{A}_{12} = \frac{21}{2500}\omega_0$, $\mathcal{A}_{20} = \omega_1$, $\mathcal{A}_{21} = \frac{119}{400}\omega_1$ and $\mathcal{A}_{22} = \frac{4641}{640000}\omega_1$, with $\omega_0 \equiv 2\sqrt{2\pi}2^{1/10}\Gamma\left(\frac{21}{10}\right) C_1 \approx 6.48562$ and $\omega_1 \equiv \sqrt{2\pi}2^{1/5}\Gamma\left(\frac{16}{5}\right) C_1^2 \approx 9.28569$. Similarly

$$\mu_4 = \sum_{k=0}^{2} \sum_{n=0}^{2} \mathcal{B}_{kn} \delta'^{\,k} a_2^n \tag{25}$$

with $\mathcal{B}_{00} = 0$, $\mathcal{B}_{01} = 4\sqrt{2\pi}$, $\mathcal{B}_{02} = \frac{1}{8}\sqrt{2\pi}$, $\mathcal{B}_{10} = \frac{56}{10}\omega_0$, $\mathcal{B}_{11} = \frac{1806}{250}\omega_0$, $\mathcal{B}_{12} = \frac{567}{12500}\omega_0$, $\mathcal{B}_{20} = \frac{77}{10}\omega_1$, $\mathcal{B}_{21} = \frac{149054}{13750}\omega_1$ and $\mathcal{B}_{22} = \frac{3484424}{5500000}\omega_1$.

Thus, Eqs.(23) and (19), together with Eqs.(24) and (25) form a closed set to find the time evolution of the temperature and of the coefficient a_2. In contrast to the case of constant restitution coefficient where $a_2 = \text{const.}$, in a gas of viscoelastic particles the time evolution of temperature is coupled to the time

evolution of the Sonine coefficient a_2. This coupling may lead to a rather peculiar time-evolution of the system.

Introducing the reduced temperature $u(t) \equiv T(t)/T_0$ we recast the set (23), (19) into the form

$$\dot{u} + \frac{1}{\tau_0} u^{\frac{8}{5}} \left(\frac{5}{3} + \frac{2}{5} a_2 + \frac{7}{500} a_2^2 \right) - \frac{1}{\tau_0} q_1 \delta \, u^{\frac{17}{10}} \left(\frac{5}{3} + \frac{119}{240} a_2 + \frac{1547}{128000} a_2^2 \right) = 0 \quad (26)$$

$$\dot{a}_2 - r_0 u^{1/2} \mu_2 \, (1 + a_2) + \frac{1}{5} r_0 u^{1/2} \mu_4 = 0 \quad (27)$$

where we introduce the characteristic time

$$\tau_0^{-1} = \frac{16}{5} q_0 \delta \cdot \tau_c(0)^{-1} = \frac{16}{5} q_0 \delta \left(4\sqrt{\pi} g_2(\sigma) \sigma^2 n \sqrt{T_0} \right)^{-1} \quad (28)$$

and define $q_0 = 2^{1/5} \Gamma(21/10) C_1/8 \approx 0.173318$, $q_1 \equiv 2^{1/10}(\omega_1/\omega_0) \approx 1.53445$, and $r_0^{-1} \equiv (24\sqrt{2\pi}/5) q_0 \delta \tau_0$. Equation (28), as shown below, describes the time evolution of the temperature. To obtain the last equations we use the expressions for μ_2, μ_4, B, and for the coefficients $\mathcal{A}$. Note that the characteristic time τ_0 is $\delta^{-1} \gg 1$ times larger than the mean collision time $\tau_c(0)$.

We will find the solution to these equations as expansions in terms of the small dissipative parameter δ ($\delta'(t) = \delta \cdot 2^{1/10} u^{1/10}(t)$):

$$u = u_0 + \delta \cdot u_1 + \delta^2 \cdot u_2 + \cdots, \qquad a_2 = a_{20} + \delta \cdot a_{21} + \delta^2 \cdot a_{22} + \cdots \quad (29)$$

Substituting Eqs. (29), (24) and (25) into Eqs. (26), (27), one can solve these equations perturbatively for each order of δ.

Keeping only linear terms with respect to δ, one can find the analytical solution (see [17] for details). This reads e.g. for the coefficient $a_2(t)$:

$$a_2(t) = \delta \cdot a_{21} = -\frac{12}{5} w(t)^{-1} \left\{ \mathrm{Li}\left[w(t)\right] - \mathrm{Li}\left[w(0)\right] \right\} \quad (30)$$

where $w(t) \equiv \exp\left[(q_0\delta)^{-1} \left(1 + t/\tau_0\right)^{1/6} \right]$ and $\mathrm{Li}(x)$ is the logarithmic integral. For $t \ll \tau_0$ Eq.(30) reduces to

$$a_2(t) = -\delta \cdot h \left(1 - e^{-8t/15\tau_c(0)} \right) \quad (31)$$

where $h \equiv 2^{1/10} \left(\mathcal{B}_{10} - 5\mathcal{A}_{10} \right) / 16\pi = 0.415964$. As it follows from Eq.(31), after a transient time of the order of few collisions per particle, i.e. for $\tau_c(0) < t \ll \tau_0$, $a_2(t)$ saturates at the "steady-state" (on the time-scale $\sim \tau_c(0)$) value $-h\,\delta = -0.415964\delta$. For $t \gg \tau_0$ it decays on a "slow" time-scale $\sim \tau_0$:

$$a_2(t) \simeq -\delta \cdot h \left(t/\tau_0 \right)^{-1/6} \quad (32)$$

and the velocity distribution tends to the Maxwellian. Linear theory gives for the temperature for $t \gg \tau_0$ [17]

$$\frac{T(t)}{T_0} = \left(1 + \frac{t}{\tau_0} \right)^{-5/3} + \left(\frac{12}{25} h + 2q_1 \right) \left(\frac{t}{\tau_0} \right)^{-11/6} \quad (33)$$

with the constants h and q_1 given above. The linear theory agrees fairly well with numerical solution (see Fig. 1 where the time dependence of a_2 is given) for small δ.

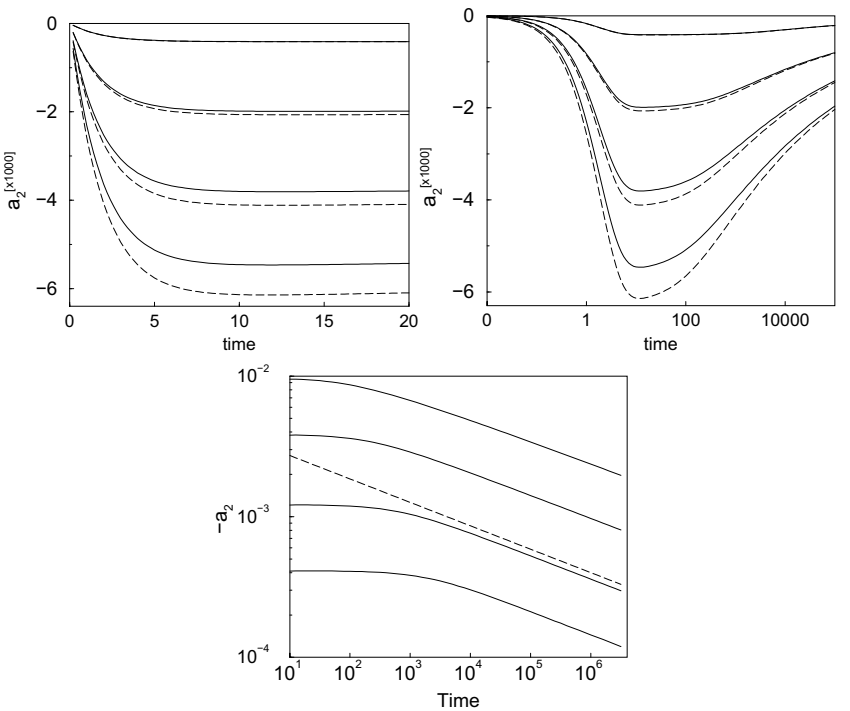

Fig. 1. Time dependence of the second coefficient of the Sonine polynomial expansion $a_2(t) \times 100$. Time is given in units of the mean collision time $\tau_c(0)$. $\delta = 0.1, 0.11, 0.12, \ldots, 0.20$ (bottom to top).

For larger values of δ the linear theory breaks down and we performed only numerical study of the equations. The results are given in Fig.2. As compared to the case of small δ, an additional intermediate regime in the time-evolution of the velocity distribution is observed. The first "fast" stage of evolution takes place, as before, on the time scale of few collisions per particle, where maximal deviation from the Maxwellian distribution is achieved (Fig.2). For $\delta \geq 0.15$ these maximal values of a_2 are positive. Then, on the second stage (intermediate regime), which continues $10 - 100$ collisions, a_2 changes its sign and reaches a maximal negative deviation. Finally, on the third, slow stage, $a_2(t)$ relaxes to zero on the slow time-scale $\sim \tau_0$, just as for small δ. In Fig.2 (left) we show the first stage of the time evolution of $a_2(t)$ for systems with large δ. At a certain value of the dissipative parameter δ the behavior changes qualitatively, i.e. the system then reveals another time scale as discussed above.

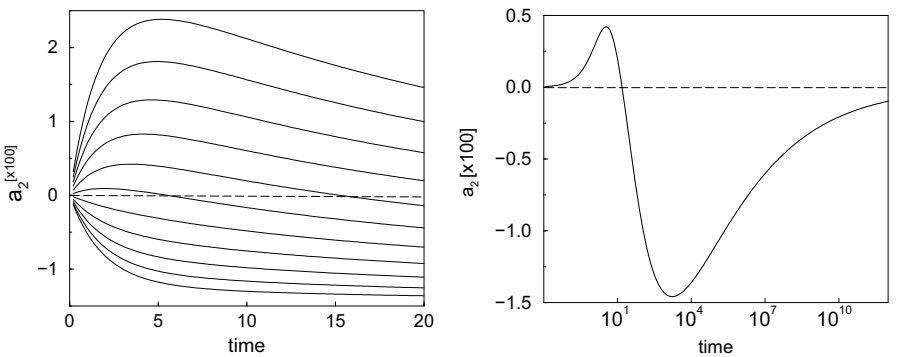

Fig. 2. The Sonine coefficient $a_2(t)$ for larger dissipation δ (numerical results). Time is given in units of mean collisional time $\tau_c(0)$. Left: $a_2(t) \times 100$ for $\delta = 0.1, 0.11, 0.12, \ldots, 0.20$ (bottom to top). Right: The plot of $a_2(t) \times 100$ for $\delta = 0.16$ over logarithmic time shows all stages of evolution discussed in the text.

5 The Age of a Granular Gas

Assume that the distribution function at time of initialization of a granular gas is known to be Maxwellian, e.g. from the nature of the physical process which gave rise to the granular gas. Then the calculation in the previous section describes the evolution of the velocity distribution function quantified by the first non-trivial Sonine coefficient a_2. We have seen that this quantity evolves characteristically in time. Assume we know an experimental method to measure the velocity distribution over a certain time interval (t_1, t_2) and, therefore, to trace the time evolution of a_2 in this interval. The time dependence of a_2 is known theoretically (at least numerically) as the solution of the set of differential equations (26,27), which depends on the parameters τ_0 and δ, i.e. $a_2 = a_2(t, \tau_0, \delta)$ (τ_0 and δ, in their turn, depend on the material parameters ρ and A). This suggests a method to *measure* the age of a granular gas, i.e., to determine the time of its initialization.

Let us explain this in more detail. If we know $a_2(t)$, $t_1 \le t \le t_2$ we can compute the values of τ_0 and δ, which completely parameterize the dependence $a_2(t)$, with an accuracy, which depends on the size of the time interval $t_2 - t_1$. Hence, using the solution (either numerical or analytical) of the set of equations (26,27) we can trace back the dependence $a_2(t)$ for times $t < t_1$. The time t_0 when the curve $a_2(t)$ cuts the abscissa corresponds to the Maxwellian distribution, i.e., it gives the initialization time and, therefore, the age of the gas. In Fig. 2 (right) we see that for larger dissipation there are two such times when $a_2\left(t_0^{(1)}\right) = a_2\left(t_0^{(2)}\right) = 0$ ($0 = t_0^{(1)} < t_0^{(2)}$) but only the earlier one, $t_0^{(1)}$, corresponds to the age (see Fig. 3). Thus, we need a method to discriminate between them: If $a_2(t)$ was positive for at least a part of the time interval (t_1, t_2), then the initialization time t_0 is uniquely determined (Fig. 3a). If the value of a_2 was negative (Fig. 3b,c) we can trace time-backwards the dependence $a_2(t)$ until at time t_0 the condition

$a_2(t_0) = 0$ is fulfilled. Then additional analysis is required: The value of δ which was already computed determines the "steady-state" value of a_2 which is reached after the quick relaxation in a time of the order of few mean-collision times. For small dissipation δ it is negative and for large dissipation it is positive. Thus, if the steady state value of a_2 is negative, the time t_0 with $a_2(t_0) = 0$ corresponds to the initialization time, i.e. $t_0 = t_0^{(1)}$. Otherwise (i.e. for positive steady-state a_2), t_0 corresponds to $t_0^{(2)} > 0$, and one has to trace $a_2(t)$ further time-backward in order to find the next time t_0 which fulfills $a_2(t_0) = 0$ to find the time of initialization.

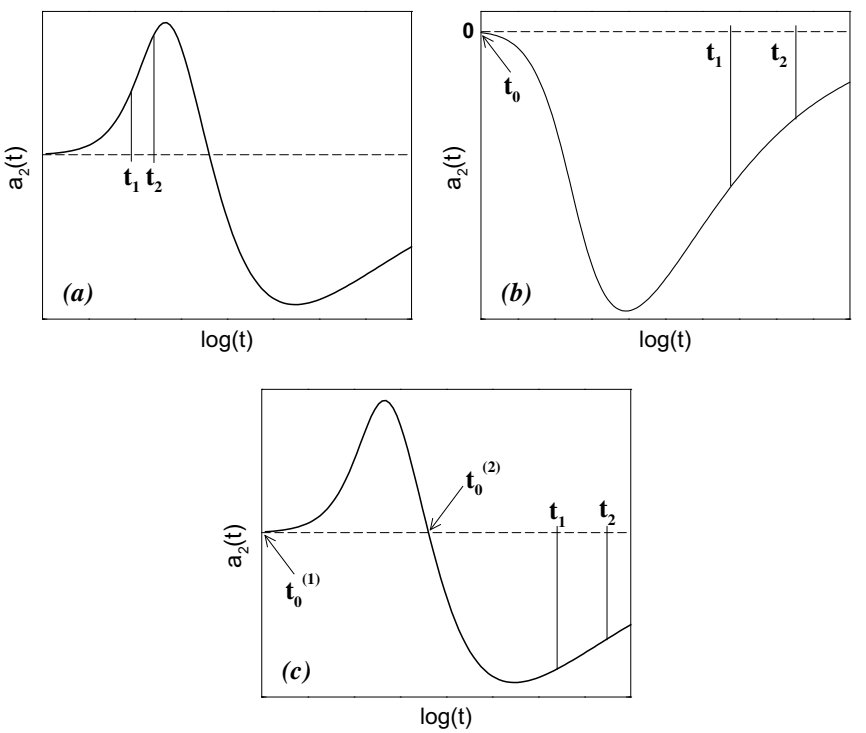

Fig. 3. Illustration of the method to compute the age of the granular gas: If the velocity distribution and, hence, $a_2(t)$ can be measured in the interval (t_1, t_2), the function $a_2(t)$ can be traced backwards in time due to the described theory. If $a_2(t) > 0$ for $t_1 < t < t_2$ (Figure (a)), the condition $a_2(t_0) = 0$ yields the age of the gas. If $a_2(t) < 0$ for $t_1 < t < t_2$ depending on material properties there may be one time $a_2(t_0) = 0$ (Figure (b)) or two times for which $a_2(t_0^{(1)}) = a_2(t_0^{(2)}) = 0$ (Figure (c)). To discriminate the cases (b) and (c) one needs further consideration (see text).

If the age of the gas has been measured according to the described method one can also calculate the initial temperature T_0, i.e. the initial energy of the gas. From δ and T_0 according to the definition of δ one can estimate the combination

of the material parameters $A\alpha^{2/5}$, and even the size of the particles σ using Eq.(28), which relates τ_0, T_0, δ and $\sigma^2 g_2(\sigma)$, provided the number density n may be measured.

Assume an astrophysical catastrophic impact took place at a certain time and produced a granular gas cloud with Maxwellian velocity distribution. If one would be able to measure the velocity distribution function in a much later time interval $t_1 \leq t \leq t_2$ following the described procedure one would be able to determine (i) the time when the impact took place, (ii) the energy of the impacting bodies (from the initial temperature), (iii) some material properties of the bodies and (iv) the grain size of the granular gas. The described analysis does not require the knowledge of the material properties of the particles.

6 Conclusion

We analyzed the time evolution of the velocity distribution function in a granular gas of viscoelastic particles in the homogeneous cooling state. The assumption of viscoelasticity is the simplest assumption for the dissipative collision of particles which is in agreement with mechanical laws. The collision of these particles is characterized by an impact-velocity dependent restitution coefficient.

For the case of a gas of particles which interact via a constant restitution coefficient its evolution is completely determined by the time dependence of the temperature. The velocity distribution function has a simple scaling form, i.e., it depends only on the reduced velocity of the particles, $\boldsymbol{c} = \boldsymbol{v}/v_0(t)$, which is just the velocity measured in units of the characteristic velocity $v_0(t)$, related to the current temperature. The scaling form, thus, persists with time.

Contrary, the velocity distribution function of a gas of viscoelastic particles does not have a simple scaling form. The deviation of the velocity distribution from the Maxwellian which is for the case $\epsilon = \text{const.}$ a function of ϵ only, i.e. time-independent, depends for a gas of viscoelastic particles explicitely on time, i.e. the velocity distribution function undergoes a time evolution. We quantify the deviation from the Maxwellian distribution by means of the first non-vanishing term of the Sonine polynomials expansion, characterized by the coefficient a_2. We assume that inelasticity of the particles is small and, hence, higher order terms may be neglected. Contrary to the case of the constant restitution coefficient, where $a_2 = \text{const.}$, for a gas of viscoelastic particles a_2 reveals a rather complex time behavior with different regimes of evolution.

The time dependence of the distribution function, quantified by $a_2(t)$, exhibits different stages of evolution, which allows to assign a Granular Gas the term "aging". The explicite time dependence implies that the process has a definite starting point of initialization, i.e. earlier times do not correspond to a physically meaningful state of the system. At first glance it might appear cumbersome that a granular gas, i.e. a cooling gas of dissipatively colliding particles, has a well defined initialization time. Indeed, if the temperature was the only characteristics of the system there would be no reason to mention some specific initial temperature, since the gas may exist at any temperature. In this case we

would not be able to conclude whether the present state of a gas is the starting point of its evolution or an intermediate one, so that its history started long ago at much higher temperature. This is the case if the particles interact via a constant restitution coefficient $\epsilon = \text{const}$. For gases of viscoelastic particles, however, the velocity distribution evolves in a way which allows to determine univocally the time-lag from the starting point and, hence, the age of the granular gas.

From the knowledge of the velocity distribution in a certain time interval $t_1 \leq t \leq t_2$ we can determine not only the age of the gas but also its initial temperature and certain material properties of the particles. This result may be useful to determine the time, the energy and other system properties of catastrophic impacts in astrophysical systems.

References

1. T. Pöschel and S. Luding (eds.), *Granular Gases*, Springer (Berlin, 2000).
2. W. Goldsmith, *Impact: The Theory and Physical Behaviour of Colliding Solids*, Edward Arnold (London, 1960); R. M. Brach, J. Appl. Mech. **56**, 133 (1989); S. Wall, W. John, H. C. Wang, and S. L. Goren, Aerosol Sci. Tech. **12**, 926 (1990); P. F. Luckham, Pow. Tech. **58**, 75 (1989).
3. F. G. Bridges, A. Hatzes, and D. N. C. Lin, Nature **309**, 333 (1984).
4. R. Ramírez, T. Pöschel, N. V. Brilliantov, and T. Schwager, Phys. Rev. E **60**, 4465 (1999).
5. N. V. Brilliantov, F. Spahn, J.-M. Hertzsch, and T. Pöschel, Phys. Rev. E **53**, 5382 (1996).
6. G. Kuwabara and K. Kono, Jpn. J. Appl. Phys. **26**, 1230 (1987).
7. T. Schwager and T. Pöschel, Phys. Rev. E **57**, 650 (1998).
8. I. Goldhirsch and G. Zanetti, Phys. Rev. Lett. **70**, 1619 (1993).
9. S. McNamara and W. R. Young, Phys. Rev. E **50**, R28 (1993).
10. T. P. C. van Noije, M. H. Ernst, R. Brito and J. A. G. Orza, Phys. Rev. Lett. **79**, 411 (1997).
11. N. V. Brilliantov and T. Pöschel, Phys. Rev. E **61**, 1716 (2000).
12. H. Hertz, J. f. reine u. angewandte Math. **92**, 156 (1882).
13. S. E. Esipov and T. Pöschel, J. Stat. Phys. **86**, 1385 (1997).
14. T. P. C. van Noije and M. H. Ernst, Physica A **251**, 266 (1998).
15. P. Resibois and M. de Leener, *Classical Kinetic Theory of Fluids* (Wiley, New York, 1977).
16. N. F. Carnahan, and K. E. Starling, J.Chem.Phys. **51**, 635 (1969).
17. N. V. Brilliantov and T. Pöschel, Phys. Rev. E **61**, 5573 (2000).
18. A. Goldshtein and M. Shapiro, J. Fluid. Mech. **282**, 75 (1995).
19. N. V. Brilliantov and T. Pöschel, Phys. Rev. E **61**, 2809 (2000).
20. T. P. C. van Noije and M. H. Ernst, Granular Matter **1**, 57 (1998).
21. N. V. Brilliantov and T. Pöschel, in [1].

Experimental and Theoretical Study of the Gravity Induced Granular Flow in Two-Dimensional Silos

Abraham Medina[1], César Treviño[2], Pedro J. Hernando[3], and Luis L. Bonilla[3]

[1] Subdirección de Exploración y Producción IMP.
A. P. 14-805, 07730 Mexico D. F. Mexico
[2] Departamento de Física, Facultad de Ciencias, UNAM.
A. P. 70-564, 04510 Mexico D. F. Mexico
[3] Universidad Carlos III de Madrid,
Avenida de la Universidad 30; 28911 Leganés, Spain

Abstract. Particle image velocimetry (PIV) has been used to experimentally study the velocity field during the discharge of a monolayer of monodisperse granular material (composed of 3 mm-diameter glass beads) in silos with different geometries. Silos were rectangular with the exit port located in the center or close to the lateral wall, and a convergent silo with angle $\theta_w = 60^\circ$. In all cases we characterized the mean velocity profiles and the fluctuation intensities. For silos with the exit port located at the center (symmetrical configurations), strong flow oscillations were detected, but not for the nonsymmetrical configuration. This last result is a strong indication that the mass flow rate is lower in the symmetrical silos. A simple theory was given for stationary flow in convergent silos and its corresponding parameters were calculated from experimental data. Numerical simulations show good agreement with measured velocity fields.

1 Introduction

The study of the gravity induced discharge of granular material is a very important field of research. Besides its important industrial and technological applications [1,2,3,4,5,6], there occur many interesting physical phenomena, such as arching, density waves, density fluctuations, size segregation, transversal oscillations and transient flow [7,8,9,10,11,12,13,14,15,16,17]. Experiments in silos filled with sand or other dry granular material show gravity-induced flows commonly exhibiting both regular and intermittent behavior in some relevant quantities, at several time scales. The mass flow rate (averaged over time) presents regular behavior: it is approximately constant during the discharge of a tall enough silo [1,2,7]. On the other hand, it has already been noticed that intermittent patterns in this flow display fractal noise in the pressure at the walls [12] and present long-range time-correlation of the number density near the exit [10,11,9]. Density fluctuations seem to occur around a mean value [3,11], which has stimulated different experimental and theoretical efforts to determine the mean velocity within the flow region [15,18,19,20,21,22]. However, very few of the theoretical studies have shown an analysis of the whole velocity field in stationary [18] or time-dependent flow [19].

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 420–435, 2001.
© Springer-Verlag Berlin Heidelberg 2001

In this work, we undertake experimental and theoretical studies of the velocity field in gravity induced granular flow in silos with simple geometries. Experimental determination of the flow uses a nonintrusive method named particle image velocimetry (PIV). This technique yields very good measurements of the flow structure (including the influence of the exit port position) for near two-dimensional (2D) silos with different configurations. Moreover, we characterize the frequencies associated with the possible transverse oscillations of the flow, the intensity (rms) of the velocity fluctuations, and the mean velocity for several sand heights measured from the bottom of the silo. These different results show very interesting and complex dynamical properties of the velocity field. To interpret the results on mean stationary flow, we study a simple fluid model consisting of continuity and momentum balance equations. The constitutive relation for the stress assumes Coulomb friction [23,24] of a slice of material with adjacent grains or with the walls [25,26], and an additional dynamic intergrain friction proportional to the square of the relative velocity [27]. The last assumption leads to a shear stress between grains proportional to the square of the velocity gradient in agreement with Bagnold's theory [28]. We obtain the coefficients in the constitutive relations and the boundary conditions from our experimental data for convergent silos and perform numerical simulations of the governing equations. Agreement between simulation and experimental results is reasonably good.

2 Experimental Set-Up

Experiments were performed in silos with glass walls and geometries as depicted in Fig. 1. Silos with vertical walls, Figures 1(a) and (b) had walls 100 cm tall, $W = 30.0$ cm width and 3.8 mm depth. They were filled up to $H = 82$ cm with monodisperse granular material composed of spherical glass-beads with mean diameter $d = 3.15 \pm 0.04$ mm, mean friction coefficient $\mu = 0.57$ and grain density $\rho_p = 2.45$ gr/cm^3. Under gravity flow, grains pass through a bottom exit with size $D = 2$ cm, located either at the center (case A), or close to the lateral wall (case B). The ratio of silo width to exit size was $W/D \simeq 15$. On the other side, we also had an experimental configuration (case C) with a convergent 2D silo having $H = 60$ cm and an opening angle $\theta_w = 60°$, Fig. 1(c). In all our experiments, the ratio of the thickness of the silo to the particle diameter is close to 1.2, which ensures assuring a monolayer of (no overlapping) grains.

The velocity field measurements were made in three different phases: (I) Flow recording, (II) digitization and (III) PIV analysis. The flow was recorded with a fast-shutter-speed (1/500 s) CCD camera with a video rate of 30 frames per second (30Hz). In the second phase (II), each recorded frame was fed to a computer and digitized as a digital image with a 256-gray scale. Finally, in phase III we used the PIV analysis software to obtain the grain displacements between two consecutive frames.

By employing this technique, the motion of a large number of small groups of grains can be determined. Measurements are obtained from each interrogation

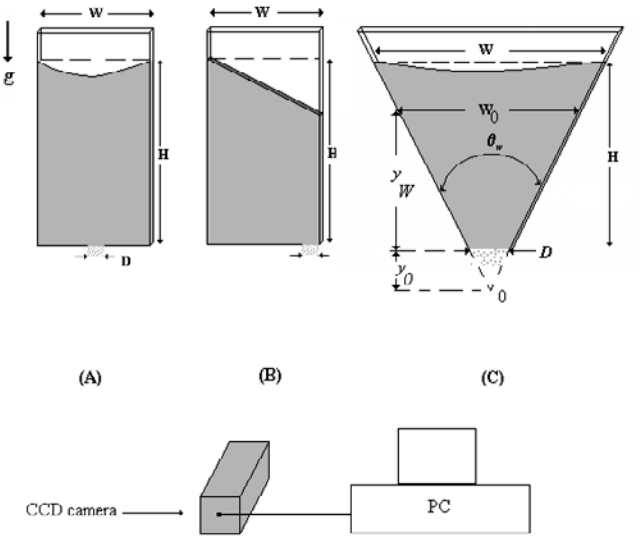

Fig. 1. Schematics for the different cases: (A) Central flow, (B) lateral flow and (C) flow in a symmetrical 2D convergent silo. In all cases $D = 2$ cm and H is the initial level. In (C) the angle of silo is θ_W, W_0 is width at the height y_W and y_0 is the length from the exit to the virtual vertex 0.

region (specifically, a small area with approximately 10 or more pairs of grains) for each two consecutive digitized images with an elapsed time, Δt, between them. Each interrogation region is located at $\mathbf{x} = (x, y)$. Vectorial displacements $\Delta\mathbf{x}$ of particle images within this region are determined through two-frame cross-correlation (see for instance [30]). Here, each two consecutive images are correlated, and the velocity vector, $\mathbf{v} = (u(t), v(t))$, is obtained at each fixed position (the center of the interrogation area) in the planar flow field as $\mathbf{v} = M\Delta\mathbf{x}(\mathbf{r})/\Delta t$. In this formula, M is the image magnification factor, $\Delta t = 1/30$ s is the time interval between images (frames), and $\mathbf{r} = (x', y')$ refers to the centroid position of the interrogation region at a given time t. Therefore, with this technique we can obtain the velocity field in the extended region covered by the image. In the experiments with the rectangular silos we used a measurement region of 30 cm height by W =30 cm width. In the convergent silo, the measurement region was $y_W \simeq 30$ cm tall (measured from the silo bottom) and W_0 =38 cm wide. In all cases, the width W or W_0 was the length of the flow region at the top of images and all dimensionless lengths will be normalized with this quantity. The number of interrogation areas, in the measurement region, along the x (transverse) and y (longitudinal) directions must be large enough to obtain a good resolution of the velocity field. As a result the PIV software produces data files containing N velocity vectors (in our case $N = 30$) at a given height, y. The number of

different heights was L (here $L = 20$), so we have $N \times L$ (600) velocity vectors in all the measurement region at each time. Finally we obtained 600 frames, which represents 20 s of the symmetric granular flow (cases A and C) and 360 frames for the nonsymmetric flow (case B). Under these conditions, errors (about 1% for velocities) primarily result from the uncertainty in determining the centroids of the particles.

3 Initial Transient Flow

Velocity vectors at different times after opening the exit port (during the initial transient) are shown in Fig. 2 for case C, the convergent silo. The other silo configurations can be treated in similar form. At the time of opening the exit hole, the flow field is confined to a region whose boundary propagates with time both upward and toward the opposite silo wall; see Fig. 2. Above that boundary, grains are closely packed and do not move. Below the boundary, grains are loosely packed, few frictional contacts occur and the grains move with a certain acceleration under the influence of gravity. Qualitative aspects of the transient initial flow have been already analyzed (see for instance ref. [13]).

Let us estimate the order of magnitude of density wave propagation in this inertia-dominated phenomenon. Dimensional arguments yield a grain velocity of $v_E \sim \sqrt{gD}$ at the exit port of the silo, provided the exit size D is large compared to particle diameter, d. Far above the exit, the characteristic velocity follows from mass conservation, $v \sim \sqrt{gD^3}/W_0$. Here $W_0 = D\,(1 + y/y_0) = 2\,(y + y_0)\,\tan(\theta_W/2)$ is the width of the flow zone located at height $y = y_W$, and y_0 is the distance from the imaginary flow vertex (denoted by 0) to the bottom exit port. At the density wave front, the grains reach the assumed velocity for fully developed flow, $v \sim g\,t_c \sim \sqrt{gD^3}/W_0$, where t_c is the characteristic time needed for a grain to reach such velocity. The characteristic time is then $t_c \sim D^{3/2} W_0^{-1} g^{-1/2}$. Using this estimation, we conclude that the density front obeys the approximate equation $dY/d\tau = 1 + Y$, where $Y = y_W/y_0$ and $\tau = d\sqrt{g/D}t/y_0$. Thus $\tau = \ln(Y + 1)$. We have used that grain size is much smaller than silo size and ignored friction. By using $\tau = \ln(Y + 1)$, $y_0 \sim 3$ cm, $d \sim 0.3$ cm and $D \sim 2$ cm, we get a time of 1 s for the flow to fully reach our measurement zone ($y_W \simeq 30$ cm). Visually, $t \sim 0.3$ s from Fig. 2 (notice that the lower part of the measurement region is 1.27 cm above the bottom exit and just 3 cm from the virtual vertex).

4 Velocity Field in Developed Flow

4.1 Transversal Oscillations

We may estimate the frequencies of the observed transversal oscillations by using the velocity estimates found in the previous Section. At the exit, $v_E \sim \sqrt{gD} \sim 40$ cm/s, and far above, at $y_W \simeq 30$ cm, $v \sim \sqrt{gD^3}/W \sim 3$ cm/s. The frequencies associated to these velocities are $f_E \sim \sqrt{gD}/D \sim 20$ Hz and $f \sim \sqrt{gD^3}/W^2 \sim$

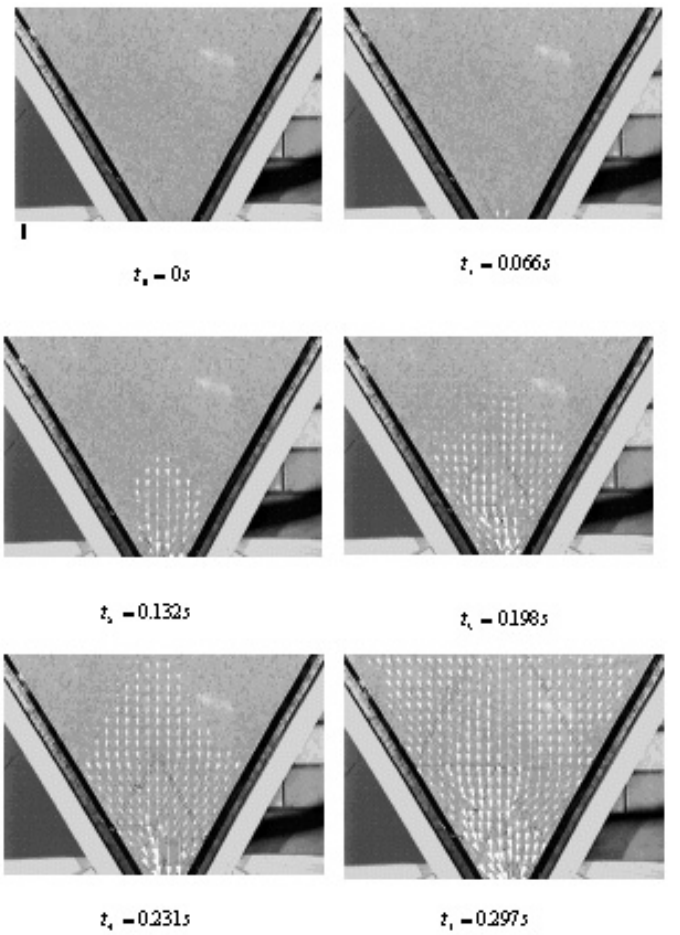

Fig. 2. Transient initial flow in the convergent silo. In this case the figures show the advance of the velocity field.

0.06 Hz, respectively. The large disparity in these frequencies means that a large number of characteristic times are involved in the flow process. Transversal oscillations of the velocity field were noted in all experimental runs, except when the exit port was located close to the lateral wall of the silo (case B) [17]. For case A, we obtained a common main frequency of flow oscillations independent of height, while for case C, location of the main frequency peak depends on height (level where the field was measured). To evaluate oscillation main frequency, we have obtained the power spectra of $G(t)$,

$$G(t) = X - < X >, \tag{1}$$

where

$$X = \frac{1}{N\overline{v}} \sum_{i=1}^{N} x_i v_i, \tag{2}$$

and N corresponds to the number ($N = 30$ in our case) of data points i at a given height and x_i is the value of x at node i measured from the left wall of the silo. Here $\overline{v}(t) = \left(\sum_{i=1}^{N} v_i\right)/N$, denotes a spatially averaged vertical velocity, while $< >$ denotes a time averaged quantity. Moreover, from Eq.(1) we found that $< G(t) >= 0$.

For case C (convergent silo), we found that the preferred oscillation frequency depens on height. Figures 3a and 3b show the power spectra at heights $y/W = 0.49$ and 0.77, respectively. In the first case, the main peak is close to $f_c \simeq 1.3$ Hz and to $f_c \simeq 0.9$ Hz in the second case. Thus the main frequency decreases with height showing the importance of the large scales induced far above the exit. These changes in the main flow direction have been detected previously during the symmetric discharge from flat-bottomed silos of an idealized granular material composed of big acryl rods of equal radius [15] or composed of equal-sized glass beads [16,17].

The characteristic oscillations of the velocity field can also be associated directly to a Strouhal number. In this case we can define a local Strouhal number as $St = f_c W/\sqrt{gD}$ where f_c is the height-dependent characteristic frequency of the flow. With this definition, the Strouhal number has the following functional form

$$St = \Phi\left(\frac{D}{W}, \frac{d}{D}\right), \tag{3}$$

which could be obtained experimentally. Despite the oscillations, the flow is symmetric, on average. This situation is visually reflected in the central depression of the free surface during flow.

4.2 RMS of Fluctuations

The velocity field $\mathbf{v}$ of a *non-steady* flow may be decomposed as the sum of a mean (time-averaged) velocity field $\langle \mathbf{v} \rangle$ and a fluctuating velocity field $\mathbf{v}'$ with $\langle \mathbf{v}' \rangle = 0$: $\mathbf{v} = \langle \mathbf{v} \rangle + \mathbf{v}'$ [31]. The velocity fluctuations intensity, (Tu, Tv), is the root mean square (rms) of $\mathbf{v}'$, normalized with the mean velocity as indicated by the following formula:

$$(Tu, Tv) = \left(\frac{\langle u'u' \rangle^{1/2}}{\langle u \rangle}, \frac{\langle v'v' \rangle^{1/2}}{\langle v \rangle}\right). \tag{4}$$

In Fig. 4 corresponding to case C, we show the plots of Tu (Fig. 4a) and Tv (Fig. 4b) as functions of the x/W for three different dimensionless heights ($y/W = 0.49$, $y/W = 0.69$, and $y/W = 0.77$) measured from the bottom. These patterns are typical of turbulent flow [31] (although their form is not the best or even the unique signature of turbulence), and their form is qualitatively

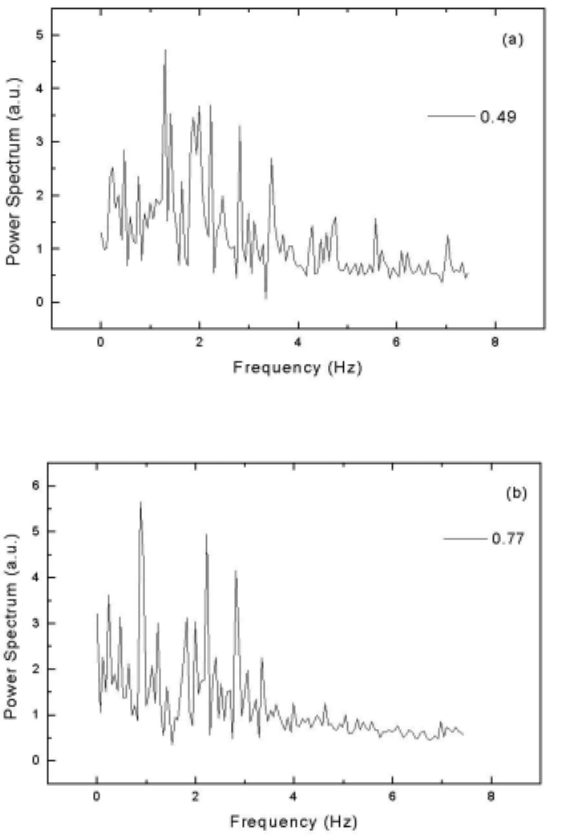

Fig. 3. Power spectrum (in arbitrary units) of the function $G(t)$ for two dimensionless heights in the convergent silo. In (a) $y/W = 0.49$ and in (b) $y/W = 0.77$. Note that the frequency in the most important peak is different in each height.

maintained for all heights. For each height, Tu has a maximum exactly at the silo center and it has two minima located symmetrically about the center. The maximum value of Tu at the center is due to the very low value of the mean velocity close to the center plane. A different situation occurs in case of Tv, where the lowest values of the root mean square of the fluctuations are smaller close to the center plane. Outside the central region, both fluctuation intensities are of the same small order of magnitude. These results are indicative of the very gentle flow in this geometry. The situation is totally different in a 2D silo with a lateral bottom exit where these fluctuations are very strong (around 10-10^3 times the velocity fluctuations for convergent silos).

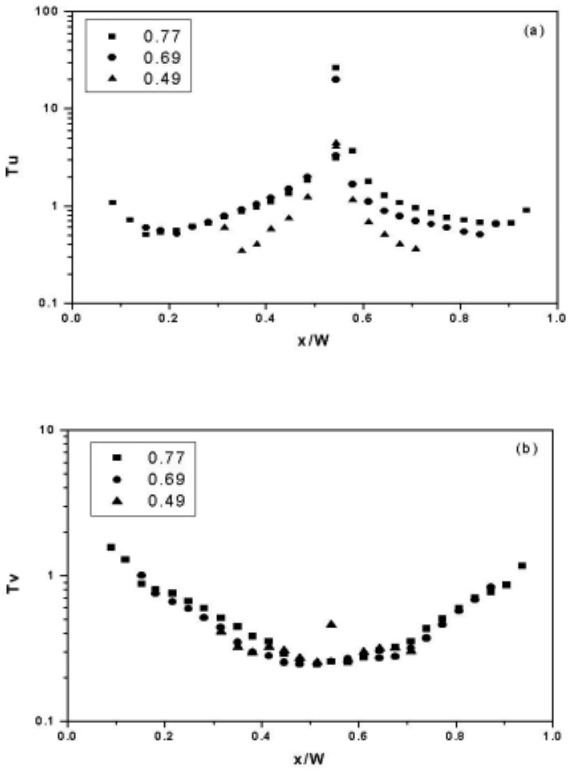

Fig. 4. RMS of the velocity fluctuations related to the mean velocity for three dimensionless heights in the convergent silo ($y/W = 0.49$, $y/W = 0.69$ and $y/W = 0.77$ from the vertex). (a) corresponds to the transversal components while (b) corresponds to the vertical components.

In Fig. 5, time-averaged velocity profiles (for near the same dimensionless height y/W) are shown for the cases A, B and C here considered. Transversal velocity is depicted in Fig. 5a, vertical velocity in Fig. 5b. The only profiles giving good agreement with the kinematic model of [16,5] were those corresponding to case A. Moreover along the vertical direction, the non-symmetric flow is larger than the symmetric ones (remember that the exit size, D, in all cases is the same). Also, the no-slip condition in cases B and C is clearly noted. This aspect will be carefully studied in future works.

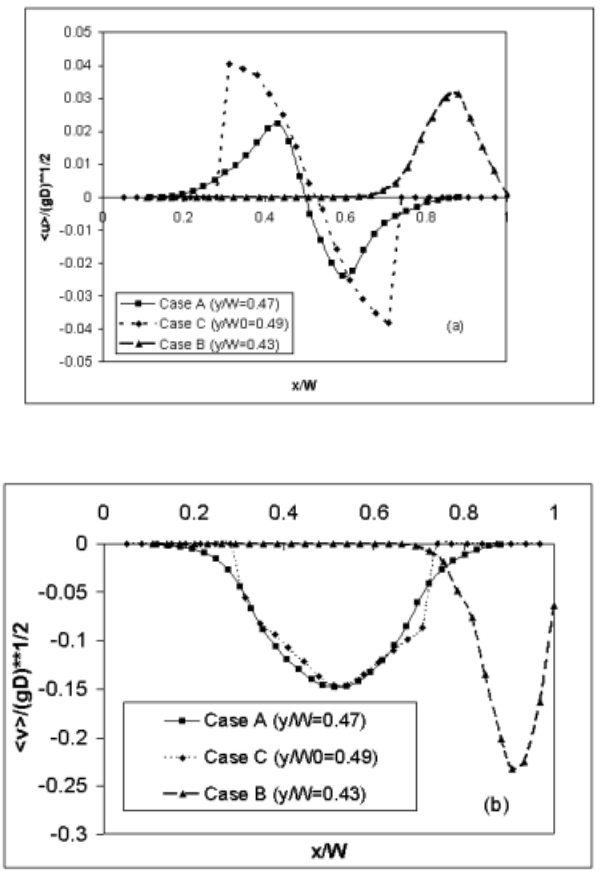

Fig. 5. Time-averaged velocity profiles for near the same dimensionless height in all cases here considered. In (a) we show the average profiles for the transversal velocity while in (b) we show the profiles for the vertical velocity.

5 Theory

Let us consider the case of a convergent silo as in Fig. 1(c). We have already seen that velocity fluctuations are much smaller in this case, so that a fluid-like theory for the mean velocity and pressure makes sense. The granular material extends from the bottom of the silo (at $y = 0$, where the half-width of the exit is $x = D/2$) to $y = H$ (where half-width is $x = W/2$). b is the depth of the hopper. Let $y = -y_0$ be the location of the vertex of the hopper and $\theta = \theta_w/2$ its half-angle; see fig. 1. Then $X(y) = (y + y_0) \tan\theta$ is the container half-width at the coordinate y. Suppose now that the motion of the sand is so slow as to ignore

transients and that the pressure p and velocity v (measured in the direction $-y$) of the material depend only on y. Obviously, H could diminish with time, so that we will fill the can with sand from above at the same rate at which the sand leaves the bottom of the container. Then H is constant and we may have a stationary 2-D flow situation. We shall consider first an idealized situation in which intergrain friction is ignored. Then trapezoidal slices of the material move compactly downward subject only to pressure forces and friction with the walls. Later on we shall consider the effect of intergrain friction.

5.1 Simple Theory Excluding Intergrain Friction

The equations of motion of the granular fluid correspond to the mass and momentum balances (in the y direction) for a trapezoidal slice of material having half-width $X(y)$ and height dy. They are $[dy/dt = -v(y)]$

$$\frac{\partial}{\partial y}(X\,v) = 0, \tag{5}$$

$$-\rho g X - \frac{\partial}{\partial y}(X\,p) + \left(1 + \frac{2X}{b}\right)\tau = \rho X v\,\frac{\partial v}{\partial y}\,, \tag{6}$$

where ρ and g are the constant density and acceleration of gravity, respectively. In (6), the forces acting on the slice are gravity, the difference between the pressures on the upper and lower sides of the slice, and the friction with the walls, τ. We shall assume that the friction of the grains with the walls has two components:

$$\tau = Kp + \alpha\rho v^2. \tag{7}$$

$2XKpdy$ is the Coulomb friction [23,24] of the slice with the wall, measured along the vertical direction. This force was used by many previous authors, in particular by Janssen [25] who analyzed the statics of sand in a cylindrical can. The other term in (7) corresponds to the resistance to the motion, and we take it to be proportional to v^2. This assumption leads to a shear stress between grains which is proportional to the square of the velocity gradient [27] in agreement with Bagnold's theory [28]; see below. The assumption that the velocity depends only on height is not too realistic. However there are two reasons to make it. Firstly, an explicit analytical solution can be found. Secondly, this theory yields appropriate boundary conditions for the velocity at the walls when we take a mean velocity field which depends on width and height.

Equation (5) yields

$$v = \frac{m}{(y + y_0)\,\tan\theta}\,, \tag{8}$$

which inserted in (6) leads to

$$\frac{\partial p}{\partial y} + \left(\frac{1 - K\cot\theta}{y + y_0} - \frac{2K}{b}\right)p = \frac{\rho\, m^2}{(y + y_0)^2\tan^2\theta}\left(\frac{1 + \alpha\cot\theta}{y + y_0} + \frac{2\alpha}{b}\right) - \rho g. \tag{9}$$

This equation is to be solved with the condition $p(H) = 0$. The additional condition that $p(0) = 0$ determines the constant m.

To analyze this problem, it is convenient to rewrite it in dimensionless units. If we use the following definitions:

$$\zeta = 1 - \frac{y}{H}\,, \qquad V = \frac{v}{\sqrt{gH}}\,, \qquad P = \frac{p}{\rho g H}\,, \tag{10}$$

Eqs. (8) and (9) become

$$V(\zeta) = \sqrt{\frac{\tan\theta}{\tan\theta+\alpha}}\,\frac{M}{1+\eta-\zeta}\,, \tag{11}$$

$$\frac{\partial P}{\partial \zeta} - \left(\frac{1-k}{1+H-\eta} - K\gamma\right) P = 1 - \frac{M^2}{(1+\eta-\zeta)^3} - \frac{\nu M^2}{(1+\eta-\zeta)^2}\,. \tag{12}$$

Here we have defined the following dimensionless parameters

$$\eta = \frac{y_0}{H}\,, \quad k = K\cot\theta, \quad \gamma = \frac{2H}{b},$$
$$M = m\sqrt{\frac{\tan\theta+\alpha}{gH^3\tan^3\theta}} \quad \nu = \frac{\alpha K\gamma\tan\theta}{K\left(\tan\theta+\alpha\right)}\,. \tag{13}$$

Eq. (12) has to be solved with the condition $P(0) = 0$ (zero pressure at the top of the silo). M is then determined by imposing that $P(1) = 0$ (zero pressure at the silo bottom). The results are:

$$P(\zeta) = (1+\eta-\zeta)^{k-1}\int_0^\zeta \frac{e^{-\gamma K\,(\zeta-z)}}{(1+\eta-z)^{k+2}} \times \left[(1+\eta-z)^3 - M^2\left[1+\nu\,(1+\eta-z)\right]\right]\,dz\,, \tag{14}$$

$$M^2 = \frac{\int_0^1 e^{\gamma K z}(1+\eta-z)^{1-k}dz}{\int_0^1 e^{\gamma K z}(1+\eta-z)^{-2-k}[1+\nu\,(1+\eta-z)]\,dz}\,. \tag{15}$$

Let us examine two limiting cases. If $\gamma \to 0$, we have:

$$M^2 = \frac{1+k}{2-k}\,(1+\eta)^{2-k}\,\eta^{1+k}\,\frac{1-\left(\frac{\eta}{1+\eta}\right)^{2-k}}{1-\left(\frac{\eta}{1+\eta}\right)^{1+k}}\,, \tag{16}$$

$$P(\zeta) = \frac{1+\eta}{2-k} f(\zeta)^{k-1}\left[1-f(\zeta)^{2-k}\right] - \frac{M^2\left[1-f(\zeta)^{1+k}\right]}{(1+k)\,(1+\eta-\zeta)^2}\,, \tag{17}$$

$$f(\zeta) = \frac{1+\eta-\zeta}{1+\eta}\,. \tag{18}$$

In the opposite case of very small silo depth compared to the height of granular material, $\gamma \to +\infty$, we have:

$$M^2 \sim \frac{\eta^3}{1+\nu\eta}\,, \qquad P(\zeta) \sim \frac{1}{\gamma K}\left[1 - \frac{\eta^3\left[1+\nu\,(1+\eta)\right]}{(1+\nu\eta)\,(1+\eta)}\right]\,. \tag{19}$$

The pressure is uniformly small and the square of the mass flow is proportional to the cubic power of the aspect ratio $\eta = y_0/H$.

5.2 Theory Including Intergrain Friction

Let us now assume that pressure and velocity depend both on the longitudinal and vertical coordinates, x and y. The velocity is supposed to reach its maximum at the center and to be a monotonically decreasing function of x at the top of the container. The balance of forces acting on a slice located between x and $x + dx$, of height dy, is

$$-\frac{\partial p}{\partial y}\,dxdy - \rho g dxdy + \left(\tau + dx\,\frac{\partial \tau}{\partial x}\right)dy - \tau dy = \rho dxdy\, v\frac{\partial v}{\partial y}\,.$$

If we now divide this expression by $dxdy$, we obtain

$$\rho v\frac{\partial v}{\partial y} = -\rho g - \frac{\partial p}{\partial y} + \frac{\partial \tau}{\partial x}\,. \tag{20}$$

The shear stress in Eq. (20) is sum of a term proportional to the pressure (corresponding to Coulomb friction of the slice with the grains adjacent to it) and a dynamical friction term proportional to $(\partial v/\partial x)^2$. To understand the latter, let us consider the dynamical friction forces on a grain on the inner boundary of the annular slice. This grain experiences two dynamical friction forces: (i) a positive force proportional to the square of its velocity relative to the faster grains at $x - d$ (d is the grain size), $F_1 = k\,(v_1 - v)^2$, and a negative force proportional to the square of its velocity relative to the slower grains at $x + d$, $F_2 = k\,(v_2 - v)^2$. The velocity of the grains at $x \pm d$ is $v_i \approx v + (-1)^i d\,\partial v/\partial y + (d/2)^2\,\partial^2 v/\partial x^2$. Then the overall dynamical friction force is $F_1 - F_2 \approx -2kd^3\partial v/\partial x\,\partial^2 v/\partial x^2 = -kd^3\partial[(\partial v/\partial x)^2]\partial y$. Thus the remaining equations of our model are the constitutive equation for τ and the balance of mass:

$$\tau = Kp + \beta\left(\frac{\partial v}{\partial x}\right)^2\,, \tag{21}$$

$$\int_0^X v(x,y)\,dx = m. \tag{22}$$

Provided $p = p(y)$ is known, (20) and (21) form a degenerate parabolic equation for v which makes sense for initial profiles such that $\partial v(x,H)/\partial x < 0$. The pressure is to be selected so that the integral constraint (22) holds for each positive value of the time-like coordinate y. Finally, the constant m is calculated by imposing that $p(0) = 0$. We should now discuss the appropriate boundary conditions holding at $x = X(y)$. Boundary conditions should be such that (6) is obtained from (20) by integrating this equation from $x = -X$ to $x = X$ under the assumption that $v = v(y)$. We get:

$$\rho X v\frac{\partial v}{\partial y} = -\rho g X - \frac{\partial (pX)}{\partial y} + \tau|_{x=X}\,.$$

A comparison of this expression with (6) shows that $\tau|_{x=X}$ should coincide with $(1+2X/b)\,\tau$, where the stress τ is given by Eq. (7). Then

$$\left(\frac{\partial v}{\partial x}\right)^2 = \frac{2X}{b\beta} Kp + \left(1+\frac{2X}{b}\right)\,\frac{\alpha\rho}{\beta}\,v^2 \tag{23}$$

is the boundary condition which holds at $x = \pm X(y)$. Provided $\alpha\rho$ and β are positive numbers, this condition is equivalent to

$$\frac{\partial v(\pm X,y)}{\partial y} = \pm\sqrt{\frac{2XK\,p(y)}{b\beta} + \left(1+\frac{2X}{b}\right)\,\frac{\alpha\rho v(\pm X,y)^2}{\beta}}. \tag{24}$$

Again, it is convenient to write these equations in nondimensional form. In addition to (10), let us define $\xi = x\,\cot\theta/H$. Then (20) - (24) yield

$$V\frac{\partial V}{\partial\zeta} = 1 - \frac{\partial P}{\partial\zeta} - B\,\frac{\partial}{\partial\xi}\left(\frac{\partial V}{\partial\xi}\right)^2, \tag{25}$$

$$\int_0^{\Xi(\zeta)} v(\xi,\zeta)d\xi = \frac{M}{\sqrt{1+\alpha\cot\theta}}, \tag{26}$$

$$\frac{\partial V(\pm\Xi,\zeta)}{\partial y} = \mp\sqrt{\frac{\gamma}{B}}\,\cot\theta\,\sqrt{K\,P(\zeta)+\alpha\,\left(\Xi+\frac{\cot\theta}{\gamma}\right)\,[V(\pm\Xi,\zeta)]^2}, \tag{27}$$

where $\Xi(\zeta) = 1+\eta-\zeta$, and $B = \beta\cot^3\theta/(\rho H^2)$.

Our mathematical problem is solving Equations (25) and (26) for $V(\xi,\zeta)$ and $P(\zeta)$ together with the boundary conditions (27). Initial conditions are $V(\xi,0)$ given and $P(0)=0$. Provided $V(\xi,0)$ is even in ξ, we will have $\partial V(0,\zeta)/\partial\xi = 0$ for any ζ. Thus we can solve our problem for $\xi > 0$ with boundary conditions $\partial V(0,\zeta)/\partial\xi = 0$ and (27) (upper sign). The parameter M should be determined so as to render $P(1)=0$. These equations contain five dimensionless parameters, α, K, B, γ and η. Out of these, γ and η are known aspect ratios, whereas the parameters α, K and B should be determined from measurements.

To determine the three unknown parameters, α, K and B, we proceed as follows. P could be found by inserting the measured $V(0,\zeta) = V_0(\zeta)$ in (25) and taking into account that $\partial V(0,\zeta)/\partial\xi = 0$. The result should be that

$$B = \frac{V_0\frac{\partial V_0}{\partial\zeta} - V\frac{\partial V}{\partial\zeta}}{\frac{\partial}{\partial\xi}\left(\frac{\partial V}{\partial\xi}\right)^2} \tag{28}$$

is an identity for all ξ and ζ. The other parameters, α and K appear in the boundary condition (27), which may be written as

$$B\tan^2\theta\,\left(\frac{\partial V(\Xi,\zeta)}{\partial y}\right)^2 = K\gamma\,P(\zeta) + \alpha\,(\gamma\Xi+\cot\theta)\,[V(\Xi,\zeta)]^2\,. \tag{29}$$

The right hand side of this Equation is linear in both K and α. We have calculated B, K and α by fitting (28) and (29) to experimental data. Our best results

are $K = 2 \times 10^{-7}$, $\alpha = 2 \times 10^{-3}$ and $\beta = 0.76$ gr mm^{-1}. Once these parameters are determined, we have solved numerically the mathematical problem above described. The results of our simulations are compared to experimental data in Figure 6. Notice that the pressure (inset) vanishes at top and bottom of the silo. Agreement between experimental and simulation data seems reasonable.

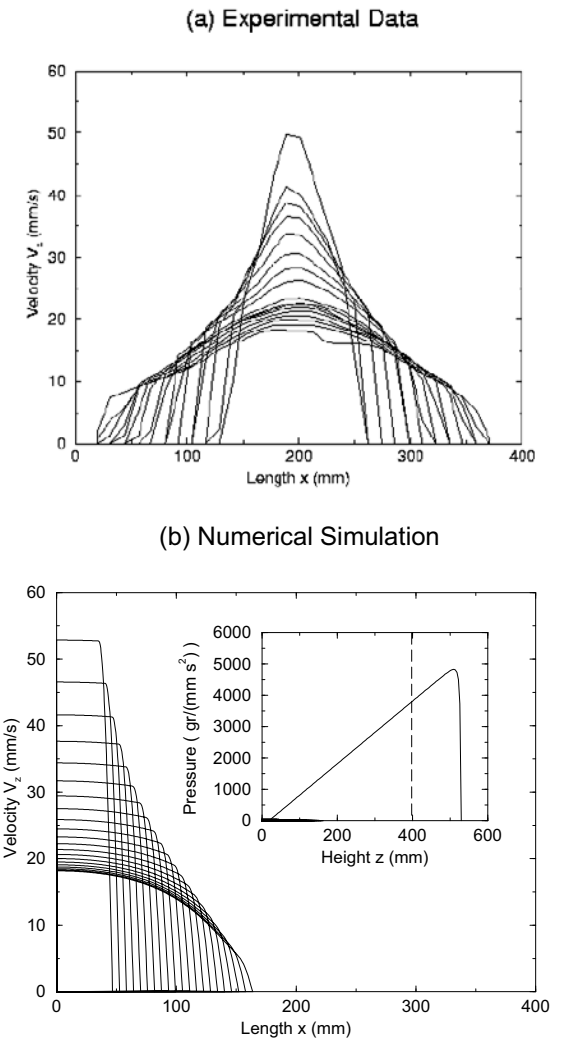

Fig. 6. Comparison between (a) experimental data and (b) simulation results for the vertical velocity as a function of width x. The inset in (b) shows the numerically calculated pressure as a function of heigth $z = H - y$, measured from the top.

6 Concluding Remarks

In this work we have used Particle Image Velocimetry (PIV) to study experimentally the steady and unsteady granular velocity field in 2D silos opened at the bottom and at the top. In each experiment, we analyzed between 360-600 images in total, covering around 12-20 seconds at 30 frames per second. We also obtained both qualitative as well as quantitative information on the flow structure for three cases. In typical experiments we obtained around 600 instantaneous velocity vector plots, covering the whole area of interest in the silos. The rich information obtained through the use of this technique was useful to obtain the initial transient flow and the fluctuation intensities given by the normalized root mean square of the fluctuating velocity fields. We also detected, in the symmetric cases, the existence of transversal flow oscillations with well characterized frequencies. Dimensional analysis showed that this mechanism is global in the sense that both the large spatial scale W and the small spatial scale D have important influences. For silos with convergent geometry, numerical simulations of a simple fluid-like granular flow theory show reasonable agreement with experimental data for the mean flow. Interesting aspects such as the no-slip condition will be studied in the future.

References

1. R.L. Brown and J.C. Richards: *Principles of Powder Mechanics*, (Oxford University Press, Oxford 1970).
2. R.M. Nedderman, U. Tuzun, S.B. Savage and G.T. Houlby: Chem. Eng. Sci. **37** 1597 (1982).
3. U. Tuzun, G.T. Houlby, R.M. Nedderman, and S.B. Savage: Chem. Eng. Sci. **37** 1691 (1982).
4. T. M. Knowlton, J. W. Carlson, G. E. Klinzing, and W.-c. Yang: Chem. Eng. Progress **32** 44 (1994).
5. R. M. Nedderman: *Statics and Kinematics of Granular Media*, (Cambridge University Press, Cambridge 1992).
6. B. J. Ennis, 'Unto dust shalt thou return'. In: *Powders and Grains 97*, ed. by R. P. Behringer and J. T. Jenkins (A. A. Balkema, Rotterdam, 1997) pp. 13–23.
7. H. M. Jaeger, S. R. Nagel, and R. P. Behringer: Rev. Mod. Phys. **68** 1259 (1996).
8. G.W. Baxter and R.P. Behringer: Phys. Rev. Lett. **62** 2825 (1989).
9. C.T. Veje and P. Dimon: Phys. Rev. E **56** 4329 (1997).
10. K.L. Schick and A.A. Verveen: Nature **251** 599 (1974).
11. J. Andrade, C. Treviño and A. Medina: Phys. Lett. A **223** 105 (1996).
12. G.W. Baxter, R. Leone and R.P. Behringer: Europhys. Lett. **21** 569 (1993).
13. P. Arteaga and U.Tuzun: Chem. Eng. Sci. **45** 205 (1990).
14. A. Samadani, A. Pradhan and A. Kudrolli: Phys. Rev. E **60** 7203 (1999).
15. H. Sakaguchi, E. Ozaki and T. Igarashi: Int. J. Mod. Phys. B **7** 1949 (1993).
16. A. Medina, J. A. Córdova, E. Luna and C. Treviño: Phys. Lett. A **250** 111 (1998).
17. A. Medina, J. Andrade, J. A. Córdova and C. Treviño: Velocity field measurements in a 2D silo with a lateral bottom exit, *Accepted in* Phys. Lett. A (2000).
18. F. Wang, C. Gardner and D. G. Schaeffer: SIAM J. Appl. Math. **52** 1076 (1992).

19. U. Haüssler and J. Eibl: Jour. Eng. Mech. **110** 957 (1984).
20. G.H. Ristow: Int. J. Mod. Phys. C **3** 1281 (1992).
21. G.A. Kohring, S. Melin, H. Puhl, H.J. Tillemans and W. Vermohlen: Computer Methods Appl. Mech. Eng. **124** 273 (1995).
22. K.D. Kafui and C. Thornton, 'Some observations on granular flow in hoppers and silos'. In: *Powders and Grains 97*, ed. by R. P. Behringer and J. T. Jenkins (A.A. Balkema, Rotterdam, 1997) pp. 511–514.
23. G. Amontons: Mémoires de l'Acad. des Sciences, Paris **206** (1699).
24. C. A. de Coulomb, *Théorie des Machines Simples*, Paris 1821.
25. H. A. Janssen: Z. Ver. Dsch. Ing. **39** 1045 (1895).
26. T. Boutreux, E. Raphaël and P.G. de Gennes: Phys. Rev. E **55** 5759 (1997).
27. A. Medina and C. Treviño: Rev. Mex. Física **42** 193 (1996).
28. R. A. Bagnold: Proc. R. Soc. London Ser. A **225**, 49 (1954), and **295**, 219 (1966).
29. M. Monti, 'Speckle photography applied to the detailed study of the mechanical behavior of granular media'. In: *Powders and Grains*. ed. by J. Biarez and R. Gourvez (A.A. Balkema, Rotterdam, 1989) p. 83-89.
30. P. Buchhave, 'Particle Image Velocimetry-Status and Trends'. In: *Experimental Heat Transfer, Fluid Mechanics and Thermodynamics 1991*. ed. by J.F. Keffer, R.K. Shah and E.N. Ganic (Elsevier, Amsterdam, 1991) p. 35.
31. W.D. McComb, *The Physics of Fluid Turbulence*, (Oxford University Press, Oxford, 1992).

Dynamics of Fracture

Farid F. Abraham

IBM Research Division, Almaden Research Center, San Jose, CA 95120.

Abstract. With the advent of scalable parallel computers, atomistic simulations are providing immediate insights into the nature of fracture dynamics by allowing us to "see" what is happening on the atomic scale. One of our most intriguing findings is a dynamic instability of the crack tip in rapid brittle fracture which prevents a crack from achieving its theoretical steady-state speed equal to the Rayleigh wave speed. Also, theory suggests that crack speeds beyond the Rayleigh speed may be forbidden. However, recent experiments, for shear dominated crack growth, report evidence to the contrary. To understand this phenomenon, we have performed molecular dynamics simulations of crack propagation along a weak interface joining to strong crystals. They show that a mode I tensile crack is indeed limited by the Rayleigh wave speed, consistent with the classical theories of fracture. However, a mode II shear dominated crack can accelerate to the Rayleigh wave speed and then nucleate an intersonic daughter crack which quickly accelerates to the longitudinal wave speed. Furthermore, crack speeds can even surpass longitudinal wave speed in materials exhibiting elastic stiffening behavior. This phenomenon is totally contradictory to predictions of classical theories.

1 Introduction

Consider a smooth, straight crack in a perfect solid. Let's stretch the solid perpendicular to the direction that the crack is pointing. If the solid material is 'brittle' and we stretch the solid enough, the crack will start moving with an increasing speed. Theory tells us that if the crack maintains its motion in the original forward direction (that is to say, it remains straight!), then the crack will eventually reach a constant velocity equal to the Rayleigh speed; that's the speed of sound on a solid surface. Why the Rayleigh speed? It's because elastic energy given to the solid by the stretch must be transferred to the crack tip in order to break atomic bonds for crack extension, and this can go only as fast as the speed of sound on the crack's surface. That appears to be a simple and reasonable description of 'how fast can cracks propagate?' However, that's not how cracks behave in most cases.

I became interested in 'how fast cracks can propagate' because cracks don't travel straight and they don't go as fast as the Rayleigh sound speed, or that's what the experimenters told me. In actual fact, there is no such thing as a perfect solid in the real world, and maybe this is the reason for the discrepancy. We know that real solids have all sorts of imperfections such as grain boundaries, impurities, and other cracks! So maybe this is the reason for crooked cracks and slower speeds. Actually, this is not the only reason, and that makes the story really interesting.

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 436–451, 2001.
© Springer-Verlag Berlin Heidelberg 2001

2 Studying Materials Failure by Computer Experiments

With the present-day supercomputers, simulation is becoming a very powerful tool for providing immediate insights into the nature of fracture dynamics. Atomistic simulations yield 'abinitio' information about crack tip formation and deformation at length scales unattainable by experimental measurement and unpredictable by continuum elasticity theory and, hence, gives additional insights into the complex mechanisms of materials failure. It becomes possible to perform an absolutely clean experiment of fracture with defect-free solids, and we are coming up with some exciting results. Like the laboratory experiments, our initial interest was to study dynamic brittle failure. We were able to follow the crack propagation over sufficient time and distance intervals so that a comparison with experiment became feasible. A detailed comparison between laboratory and computer experiments have shown commonality in behavior at the macroscopic scale. Using our 'computational microscope,' we can see what is happening at the atomic scale.

Our simulation tool is computational molecular dynamics (Abraham, 1986), and it is very easy to describe. Molecular dynamics predicts the motion of a given number of atoms governed by their mutual interatomic interaction, and it requires the numerical integration of the equations of motion, 'force equals mass times acceleration or F = ma.' We learn in beginning physics that the dynamics of two atoms can be solved exactly. Beyond two atoms, this is impossible except for a few very special cases, and we must resort to numerical methods. A simulation study is defined by a model created to incorporate the important features of the physical system of interest. These features may be external forces, initial conditions, boundary conditions, and the choice of the interatomic force law. My choice of simple interatomic force laws is dictated by my intent to investigate the generic features of a particular many-body problem common to a large class of real physical systems and not governed by the particular complexities of a unique molecular interaction. It is very important to emphasize this conscious choice since it is not uncommon to hear others object since I am not studying 'real' materials.

The simple interatomic potential may be thought of as a 'model potential,' and my model potential is normally the Lennard-Jones 12:6 potential. My 'model material' is typically a two-dimensional atomic crystal but not always. Gordon (1984) comments on Griffith's desire to have a simpler experimental material that would have an uncomplicated brittle fracture. He writes, 'In those days, models were all very well in the wind tunnel for aerodynamic experiments but, damn it, who ever heard of a model material?'

When I started out in the mid 1960s, a few hundred atoms could be treated. In 1984, we reached 100,000 atoms. Before that time, computational scientists were concerned that the speed of scientific computers could not go much beyond 4 Gigaflops, or 4 billion arithmetic operations per second and that this plateau would be reached by the year 2000! That is history. With the concept of concurrent computing, a modern parallel computer is made up of several (tens, hundreds or thousands) small computers working simultaneously on different

portions of the same problem and sharing information by 'communicating' with one another. The communication is done through 'message passing' procedures. The present record is well over a few Teraflops for optimized performance, and we recently simulated 1,000,000,000 atoms in a fracture study at the San Diego Supercomputer Center. Moore's Law states that the speeds of computers double every one and one-half years. For 35 years, that translates into a computer speed increase of ten million. This is exactly the increase in the number of atoms that I could simulate over the last 35 years.

Why do we refer to 'experiments' in the title of this section? In a certain sense, one can view the simulation methodology as giving experimental information from a model, idealized universe. This approach has a mixture of theory and experiment, and it has been labeled 'the third kind of science.' By computationally solving the theoretical models in detail, a wealth of information about a physical phenomenon is produced that can be viewed as experimental data. Like any other experimentalist, we have to ask the right questions and have to measure the right things.

3 The Dynamic Instability

Continuum fracture theory typically assumes that cracks are smooth. For dynamic cracks, it predicts that they accelerate to a limiting speed equal to the Rayleigh speed of the material. However, experiment tells us that in a common fracture sequence in brittle materials, an initially smooth and mirror-like fracturing face begins to appear misty beyond a speed of about one-third the Rayleigh speed and then evolves into a rough, hackeled region as the crack accelerates to a limiting velocity of about six-tenths the Rayleigh speed. The origin of the disagreement between theory and experiment suggests questions of microscopic impurities or imperfections in the material, experimental uncertainties, or failure of theory. All of these features are unexplained using continuum theory, and experiment suggesting a crack instability beyond one-third of the Rayleigh wave speed perked my interest in dynamic brittle fracture (Fineberg, 1991, 1992).

In order to get started, I decided to simulate the brittle fracture of one of the simplest atomic system conceivable: a perfect two-dimensional (2D) solid composed of a sheet of atoms with their motion restricted to the plane. Why two dimensions? First, it is less demanding computationally than three dimensions. Second, for small deformations the triangular lattice is elastically isotropic, and third, the two-dimensional lattice of atoms interacting with a model potential was believed to be brittle. My chosen system size was as large as two million atoms, lots of atoms in 1993. Why two million atoms? A 2D wafer of one million atoms has one thousand atoms on a side, which I found large enough to simulate the early stages of fracture dynamics. The slab length was long enough for the crack to reach its maximum speed before the elastic stress waves produced by the advancing fracture reflected off the wafer's edges and return to the fracture point. Size demands more atoms, and more atoms demand more computing,

not only because of the increased size but also because the natural time scale expands; that is, it takes more time for a crack to traverse the larger system.

The atomic slab was initialized at a very low temperature, and a notch of specified size was cut midway on a slab side. An outward strain rate was imposed on the outer most columns of atoms defining the opposing perpendicular sides of the slab. This means that we are pulling on the opposing sides parallel to the notch so that the slab is being stretched a constant percentage per unit time. An increasing lateral strain occurs in the solid slab as time progresses. At a sufficiently large strain, depending upon the choice of notch length and other things, the crack begins to move from the notch tip, and the simulation is continued until the growing crack has traversed the total length of the slab. So, what did we expect to see? Our special solid has perfect atomic packing with no crystal imperfections except for the created notch or crack. So we expected to see perfect crack motion predicted by theory; that is, with sufficient strain, the crack would accelerate to the Rayleigh wave speed with mirror surfaces being created as it passed through the solid. Were we surprized!

Figure (1) summarizes the important findings of our typical 2D fracture simulation. We see that the brittle crack initially propagates in a straight line and leaves mirror-like cleaved surfaces. However, at the crack speed of one-third of the Rayleigh wave speed (second image), the crack begins to roughen, and then to oscillate back and forth (third image) achieving a forward speed equal to approximately two-third of the Rayleigh speed. Similarly, for speeds less than one-third of the Rayleigh speed, the acceleration of the crack tip is quite smooth; however, the instantaneous tip speed becomes very erratic after reaching one-third of the Rayleigh speed. All of these features are in agreement with the laboratory findings but not with our expectation that theory would describe our idealized experiment. We didn't do just one simulation. We did many. We changed the crack length; cracks with a longer length begin moving at a lower strain and accelerate slower. We found the same thing. We changed the interatomic potential (or how the atoms interact); this changes the type of material and its sound speeds. We found the same thing. We tried bigger atomic slab sizes, different strain rates, and lots of other things. We found the same thing. So we concluded that this transition from smooth, mirror-like motion to rough, erratic motion is an instability in the dynamics of the crack. It has nothing to do with material imperfections; there weren't any in our perfect solid. It has nothing to do with a particular material; we tried different interatomic potentials. It has to do with the crack dynamics. The dynamics has an instability in the straight-line motion at one-third of the Rayleigh wave speed! This is a very important finding.

Dislocations appear after the onset of the instability and are apparent in the gray-scale pictures. They appear as dark vertical dashes in the late-time snapshot. By examination, we observe that the transverse dislocations go out some distance, then return to the fracture surface where they disappear. The spacings between these dislocations are quite regular. We can understand their origin as a transverse slippage between two neighboring rows of atoms arising from a growing shear stress at the crack tip as the ever-increasing cascade of broken

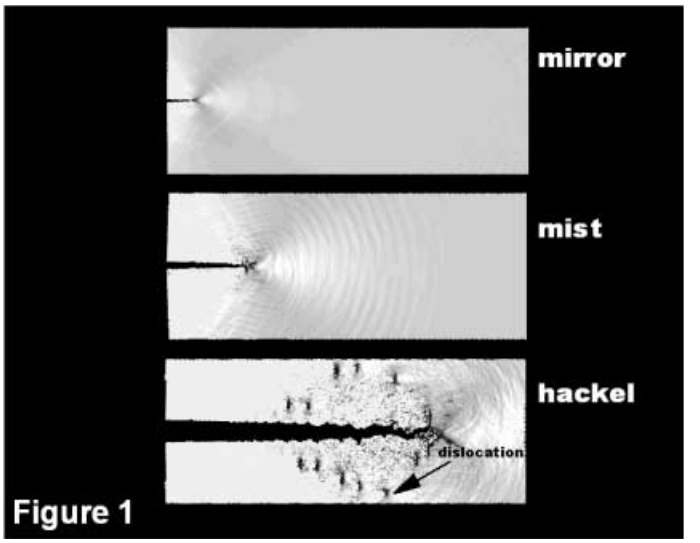

Fig. 1. The dynamic regimes of brittle fracture. The brittle crack initially propagates in a straight line and leaves mirror cleaved surfaces. At the crack speed of one-third of the Rayleigh wave speed (the second image), the crack begins to roughen, and then to oscillate back and forth (the third image) achieving a forward speed equal to two-thirds of the Rayleigh speed.

bonds in the forward direction allows ever increasing severed rows of atoms to want to relax laterally. This buildup of severed atomic rows will eventually be sufficiently large enough to overcome the barrier to slippage. The front to slippage will manifest itself as a dislocation. Of course, this is a repeating process, hence there is a continual creation of transverse dislocations. Also, the return of dislocations to the crack surface and the healing of the surface is a consequence of neighboring bands of matter bounded by slip planes relaxing to equilibrium. The dislocations disappear with no evidence that they exist during the fracture process. Examining the cleaved solid, one would erroneously conclude that dislocations are not present in the brittle fracture of our two-dimensional solid.

From a simple geometrical consideration, the oscillating zigzag motion of the crack tip explains why the 'apparent forward' crack speed is slower than theoretical prediction. It is not quite complete, but it gets it down to approximately eight-tenth of the Rayleigh speed. There are delays in the zigzagging of the crack that must be accounted for and which further reduces the speed estimated by simple geometry. If we could prevent the crack from zigzagging, its forward motion would be much faster and maybe equal to the Rayleigh speed. The atomic simulations are discussed in Abraham et. al. (1994, 1997).

4 Transonic Crack Motion

So far, I have discussed mode I (tensile) loading of a simple two-dimensional crystal with all of the atoms interacting with the same model interatomic potential. However, there are other ways to break a crystal, and the crystal structure may not be so ideal. For example, there is mode II shear loading. First, lets consider what theory tells us about the different loading.

It is widely believed that a brittle crack can not propagate faster than the Rayleigh wave speed. The origin for this belief stems from the predictions of continuum mechanics; in particular, the dynamic elastic solutions of the near-tip stress fields and energy release rates for various velocity regimes and different types of external loading (Freund, 1986; Broberg, 1999). For a mode I (tensile) crack, both the energy release rate and the stress singularity vanish for all crack velocities in excess of the Rayleigh wave speed. The implication is that there is no possibility for a mode I crack to propagate at a velocity greater than the Rayleigh wave speed. A mode II (shear) crack behaves similarly to a mode I crack in the subsonic velocity range; i.e., the energy release rate monotonically decreases to zero at the Rayleigh wave speed and remains zero between the Rayleigh and shear wave speeds. However, the predictions for the two loading modes differ for crack velocities greater than the transverse wave speed. The energy release rate is zero for a mode I crack but is positive for a mode II crack over the entire range of intersonic velocities, with a maximum value at the radiation-free speed equal to the square root of two times the transverse sound speed. What has been concluded from these theoretical solutions? First, a mode I crack's limiting speed is the Rayleigh speed. Second, a mode II crack's limiting speed is also the Rayleigh speed because the 'forbidden velocity zone' between the Rayleigh and shear wave speeds acts as an impenetrable barrier for the shear crack to go beyond the Rayleigh wave speed. The forbidden regimes for crack propagation are summarized in Figure 2 for the different applied loads.

The first direct experimental observation of cracks faster than the transverse wave speed has been reported by Rosakis et al. (1999). They investigated shear dominated crack growth along weak planes in a brittle polyester resin under far field asymmetrical dynamic loading. The late arrival of an experiment on intersonic fracture is due, in part, to the fact that a crack in elastic homogeneous and isotropic solids has a zigzag crack path once the crack tip velocity exceeds the dynamic instability speed of one-third of the Rayleigh wave speed. Therefore, a wavy crack instability occurs at low crack velocities and prevents an exploration of the full range of possible velocities. By introducing a weak path (a layer of lower toughness), the crack growth is confined to this path.

We have performed two-dimensional molecular dynamics simulations of crack propagation along a weak interface joining two harmonic crystals (see Figure 3). We assume that the interatomic forces are harmonic except for those pairs of atoms with a separation cutting the centerline of the simulation slab. For these pairs, the interatomic potential is taken to be our simple model potential that allows the atomic bond to break. Figure 4 shows the dependence of the interatomic potential energy on atomic separation. To study a shear dominated

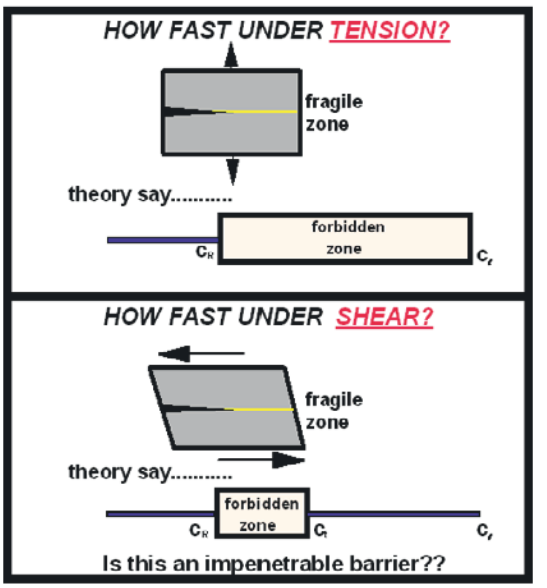

Fig. 2. The forbidden regimes for crack propagation with the different applied loads.

crack, the applied shear strain rate is five times greater than the applied tensile strain rate. The top boundary of the slab is moving up and to the right and the bottom boundary of the slab is moving down and to the left. Linear velocity gradients are established across the slab, and an increasing strain occurs in the solid slab. This leads to eventual failure of the material at the slit tip. The applied strain rates remain constant during the simulation, and the simulation is continued until the growing crack has traversed the total length of the slab. We also conducted simulations of a mode I crack with the same geometrical setup where only an opening strain rate is imposed. The results are discussed below.

The histories of crack velocity under mode I and mode II loading are easily described. The mode I crack quickly approaches a constant velocity equal to the Rayleigh wave speed of the harmonic crystal showing that the crack velocity is indeed limited by the Rayleigh wave speed and consistent with the classical theories of fracture. The mode II crack also quickly approaches a constant velocity equal to the Rayleigh wave speed. After propagating at this constant velocity for a short while, the crack tip jumps to a higher constant velocity calculated to be the longitudinal sound speed of the harmonic solid. These results are consistent with the Mach cone angles and the expanding halo configuration shown in Figure 5. The mode II crack thus exceeded the Rayleigh wave speed and contradicts the classical theories of fracture. Figure 5 shows several snap shots of a mode II crack accelerating to the longitudinal sound speed. Initially, the mode II crack is subsonic with no shock waves emanating from the crack tip, as shown in Figure 5(a). As the crack jumps over the forbidden velocity zone and attains the longi-

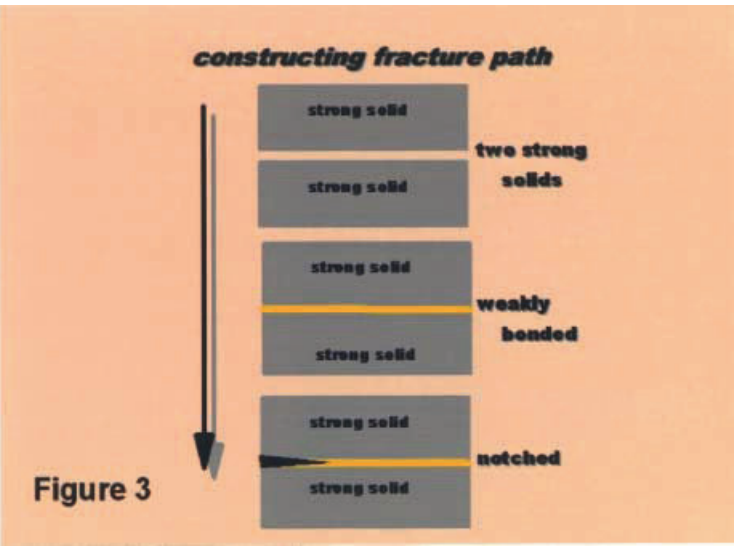

Fig. 3. Construction of the two dimensional slab that imposes crack propagation along a weak interface joining two strong crystals. The interatomic forces are harmonic except for the fragile zone defined by those pairs of atoms with a separation cutting the vertical centerline of the simulation slab. For these fragile pairs, the interatomic potential is taken to be our simple model potential (the Lennard-Jones potential) that allows the atomic bonds to break.

tudinal wave speed, a pair of shock fronts is observed as Mach cones emanating from the crack tip, as shown in Figures (5b-c). These shock fronts correspond to transverse sound waves generated at the moving crack. Note that the crack sits at the top of a circular halo that corresponds to longitudinal sound waves emitted from the crack tip. This configuration shows that the crack velocity is equal to the longitudinal wave speed.

The mechanism for the mode II crack 'jumping' over the forbidden velocity zone is clearly shown in Figure 6 where a series of color maps of the off-diagonal shear stress component is used to reveal the details of this process. This transition occurs by the nucleation of an intersonic daughter crack ahead of the mother crack traveling at the Rayleigh wave speed. As the mother crack approaches the critical state of nucleation, the crack tip region is asymmetrically distorted with a bulge on the right side of the crack face, as shown in Figures 5(a). The shear stress distribution near the crack tip is also highly asymmetric. We observe significantly less crack opening at the tip of the daughter crack compared with that of the mother crack, indicating that the daughter crack is more dominantly mode II even though the mother crack has a mode I component. This is consistent with our observation that the mode I crack tends to be limited by the Rayleigh wave speed. Figures 5(b-d) show the detailed process of the birth of the intersonic daughter crack. It is seen that a sharp intersonic crack is nucleated at a small

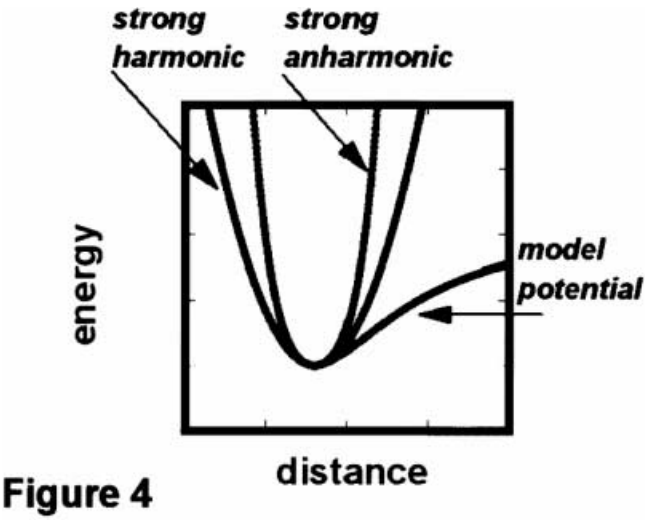

Fig. 4. The dependence of the interatomic potential energy on atomic separation for various force laws used in the simulations discussed in this paper.

distance ahead of the mother crack. The angle of the Mach cone shows that the velocity of the daughter crack is consistent with the longitudinal wave speed. As the daughter crack moves ahead, the mother crack can still be seen as a surface bulge that trails behind at the Rayleigh wave speed.

Our simulations demonstrate intersonic crack propagation and the existence of a 'mother-daughter' crack mechanism for a subsonic shear crack to jump over the forbidden velocity zone. This mechanism is reminiscent of a similar mechanism based on continuum theories (Burridge, 1973; Andrews, 1976), although the continuum description can not provide an abinitio description for crack formation and the details of crack tip distortion are not consistent with the continuum solutions. The birth of the daughter crack can not be characterized by a critical energy release rate or a critical stress intensity factor near the mother crack because both these quantities vanish at the Rayleigh wave speed. It seems that the only possible mechanism by which the daughter crack can be nucleated is by the finite stress peak ahead of the mother crack, and along the weak bonding line, as measured in the stress field and discussed by Burridge (1973). The critical condition for the nucleation of the daughter crack has been investigated using dynamic elasticity equations. The elasticity analysis shows that, while the initiation of the mother crack is governed by the Griffith criterion of fracture, the daughter crack is nucleated by cohesive failure of the weak interface ahead of the crack tip. Comparison of the relevant length and time scales predicted by the elasticity analysis with the results from molecular dynamics simulation is excellent (Gao et al., 2000).

We have also made further extensions to this study. For example, we have set the applied strain rates equal to zero at the time of nucleation of the intersonic

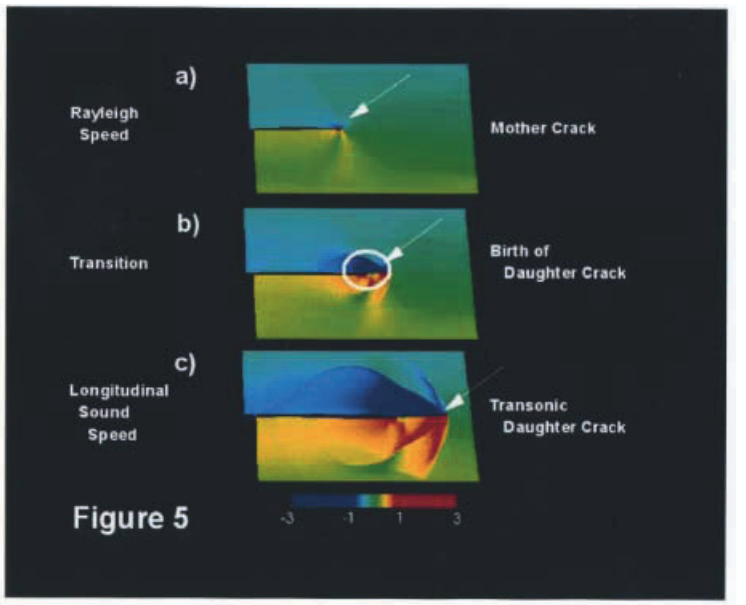

Fig. 5. Transition of a subsonic crack to the longitudinal sound speed. The figures represent a progression in time from the top to bottom. (a) A mode II shear crack travels at the Rayleigh wave speed; (b) The early-time occurrence of shock waves as the crack attains the longitudinal wave speed; (c) The late-time propagating crack traveling at the longitudinal wave speed. . The colorbar shows the color map for the local atom speed in the direction perpendicular to the moving crack. The slab dimensions are approximately 800 by 1400 in reduced units, the length scale being the atomic distance when the model interatomic potential crosses over from being repulsive to being attractive.

crack and still observed a constant propagation velocity of the crack at the longitudinal wave speed throughout its passage to complete cleavage. However, the acoustical disturbances are much less explosive. If the slab is allowed to spring back by completely releasing the applied loads as soon as the daughter crack is nucleated, the crack speed drops to the 'square-root of two times the transverse wave speed' and propagates at this speed until complete cleavage occurs. This loading process seems to closely resemble the dynamic impact loading conditions in the experiments of Rosakis et al. (1999), and our results are entirely consistent with their corresponding experimental observations. The special speed corresponds to the radiation-free speed at which the maximum intersonic energy release rate occurs. The atomic simulations for transonic crack propagation are discussed in Abraham and Gao (2000).

Can cracks go faster than the longitudinal wave speed? I discovered the answer to be yes quite accidentally. In my first attempt to simulate transonic crack propagation, I considered a spring potential adopting the repulsive part of the model potential both for stretch and for compression. In this case, the

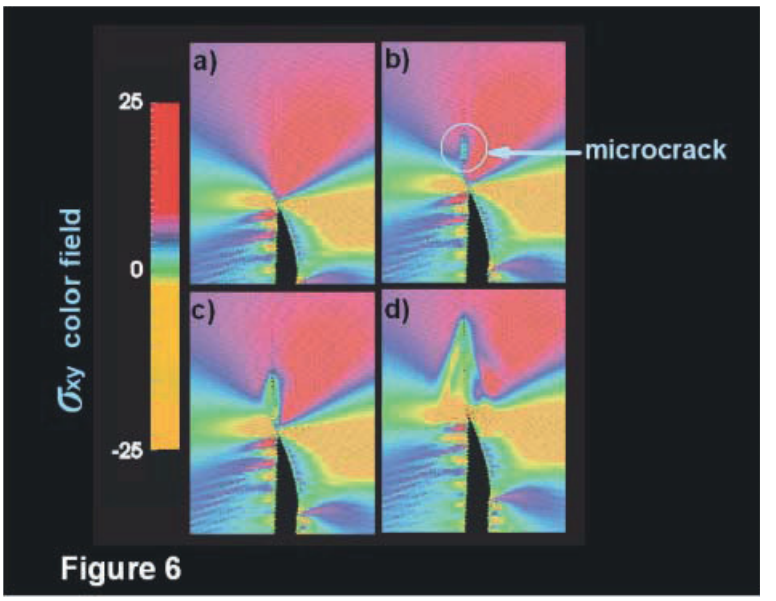

Fig. 6. The nucleation of intersonic daughter crack at the mother crack. The figures represent a progression in time from left to right and from top to bottom which are referenced as (a) to (d). (a) The approach of the critical state for the mother crack. Note the asymmetrically distorted crack tip region; (b) The birth of the intersonic crack. A very sharp slit is born ahead of the tip of the mother crack; (c, d) The daughter crack joins the mother crack and quickly approaches the longitudinal sound speed. The colorbar shows the color map for the shear stress xy component. The bulged mother crack is still propagating at the Rayleigh wave speed. The dimensions of the solid region shown are approximately 90 by 120 (in reduced units).

solid stiffens with strain, as opposed to the standard model potential for which the solid softens with stretch. In Figure 7, we see two shock cones emanating from the moving crack tip. The crack is 'supersonic' and moving 25wave speed. This phenomenon is totally contradictory to predictions of classical theories of fracture.

The nonlinear properties of the material may provide a possible explanation of this phenomenon. Gao (1993) has shown that such 'hyperelastic' softening of solids leads to a significant reduction in local wave speeds near a crack tip. We expect that the hyperelastic stiffening of solids will induce significantly higher values of wave speeds near the crack tip. Local wave speeds near the crack tip region have increased significantly beyond the longitudinal wave speed of the harmonic crystal because of the elastic stiffening in this region. As a result, the limiting fracture speed at the crack tip is no longer determined by the harmonic properties of a solid, but rather by its hyperelastic properties. When the elastic disturbance goes beyond the highly deformed tip region, it is traveling supersonically with respect to the harmonic region. Hence, two Mach cones appear.

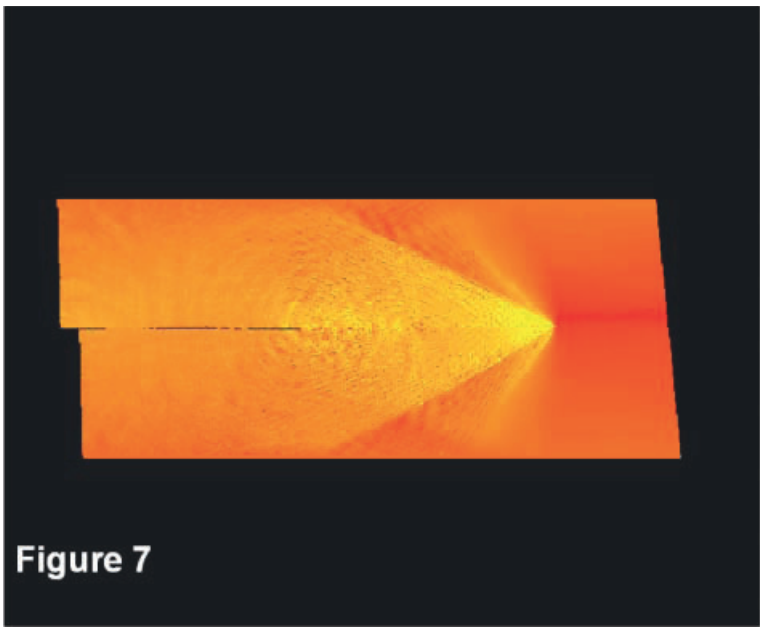

Fig. 7. A supersonic crack in an elastic-stiffening, two-dimensional solid. The slab dimensions are approximately 700 by 1600 (in reduced units).

5 A Dynamic Brittle-to-Ductile Transition

We have discussed many features of failure associated with rapid brittle fracture. Brittle fracture certainly is not the sole mechanism for materials failure when things are hit hard and fast; otherwise our world would be quite fragile. Basically there are two generic classes describing materials failure; a material is brittle or it is ductile. In the first case, chemical bonds are broken, and such a failure is easily recognized when you see glass shatter. For ductile failure, such a catastrophic event does not occur. Tough materials like metals do not shatter; they bend because plastic deformation occurs by the motion of rows of atoms sliding over one another on preferred slip-planes (dislocations) in contrast to bond-breaking. That is why car fenders are not made of brittle materials like glass. These two general classes of failure do not depend on the details of the interatomic interactions, but they do depend on the atomic packing of the three-dimensional solid. Of course, the interatomic potential dictates the packing.

We also recognize that there are a limited number of atomic packings in nature. Glasses do not have extended crystallinity because atoms are packed randomly. They have no slip-planes and therefore can not exhibit ductility arising from atomic planes slippage. With this picture, glasses should exhibit brittle failure. Three dimensional crystals do not have the isotropy of glasses. In a certain sense, one may view the crystalline solid as a defected solid because of the loss of perfect isotropy, the defect being multiplanar and infinite in extent! Because of

the breaking of the perfect isotropic symmetry, slip-planes exist, dislocations are possible, and ductility may win out over brittle failure. The face-centered-cubic (FCC) packing is known to have a strong propensity toward ductility. Crystallinity does not guarantee ductility; strong covalent-bonded solids tend to be brittle.

Our first attempt to go to three dimensions was a simulation of an fcc solid with a crack and with mode I loading (Abraham, 1997 and Abraham et al., 1997). With this atomic packing, one would expect that the solid would eventually undergo ductile failure when the uniaxial stretch became sufficient by crack blunting and no crack extension. On the contrary! The onset of failure is characterized by classic brittle fracture. There is perfect cleavage and crack extension from the accelerating crack up to a point, that point being reached when the crack velocity equals one-third of the Rayleigh wave speed (Figure 8). Then something quite dramatic happens. The crack stops, and a 'flower of loop dislocations' blossom from the crack tip. This is shown in Figure (9). The atomic system is a three dimensional slab having a total of more than one hundred thousand atoms. We show only those atoms that have a cohesive energy less than 97 percent of the ideal bulk value, reducing the number of atoms seen by approximately two orders of magnitude. The elastic energy continues to dissipate through the continued creation and motion of the dislocations (Abraham, 1997). Of course, 'one-third of the Rayleigh wave speed' is, by now, a very familiar speed in crack dynamics. We conclude that the early brittle fracture dynamics of the fcc solid is an extremely fragile phenomenon and that the onset of the instability nucleates the ductile failure mechanism of crack arrest, dislocation creation and propagation and associated crack blunting. Uniaxial stretching is applied in the [110] direction for this simulation, and we have argued that it is the hyperelasticity associated with the [110] direction that allows the initial failure to be brittle (Abraham, 1997). We have observed a 'dynamic brittle-to-ductile transition. You might say that we have gone full circle concerning 'how fast can cracks propagate.' [t]

6 Epilogue

I will discuss new directions for the computer simulation of fracture mechanics. The traditional approach for describing materials phenomena that span a range of length and time scales takes advantage of a hierarchy of physical description. A description of phenomena on a given length scale includes materials specific constants that can be calculated from a theory that describes shorter length scale phenomena; the constants entering this theory can in turn be calculated from a still shorter length scale theory, and so on. For most phenomena of interest to materials science, this sequential coupling of length scales finds its practical end at the level of ab initio theories of electronic structure. The pioneering work of Clementi (1988) provides a lucid example. They used quantum mechanical methods to evaluate the interactions of several water molecules. From this data base, they created an empirical potential for use in atomistic simulations, and

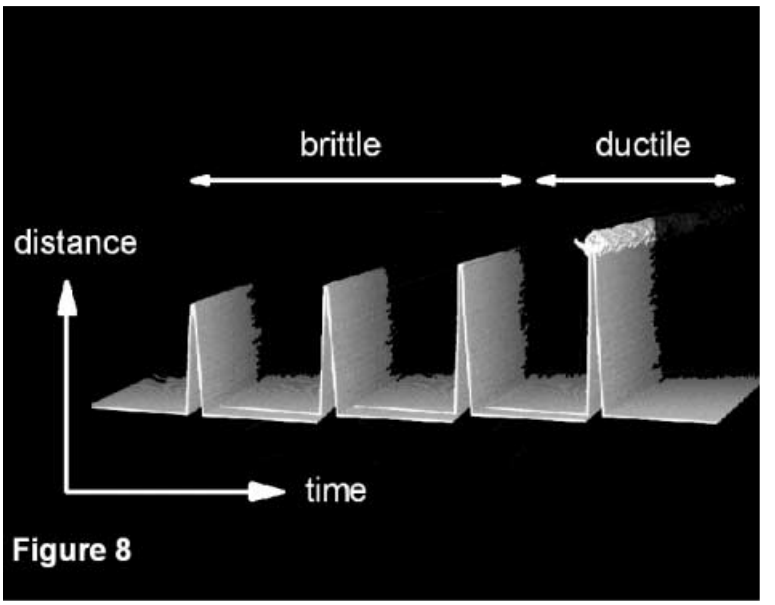

Fig. 8. Early-time sequence of the propagating crack in an FCC solid. Overlapping partial landscapes of the growing crack surface due to brittle fracture are shown in the first three images and the subsequent appearance of dislocations in the fourth image after the transition to ductility. Only atoms with a potential energy less than 97 percent of the bulk value are displayed, resulting in the selected visualization of atoms neighboring surfaces and dislocations. Approximately a third of the crack's depth is shown.

evaluated the viscosity of water. Using this computed viscosity, they performed a fluid dynamics simulation to predict the tidal circulation in Buzzard's Bay. This is a powerful example of the sequential coupling of length and time scales: a series of calculations is used as input to the next up the length/time scale hierarchy.

There are many phenomena, however, where physics on short and long length scales interact dynamically; and these can not be adequately described with a sequential coupling of length scales. Dynamic fracture is a very good example. Energy from macroscopic-scale elastic fields is concentrated on the Angstrom scale of the electrons that participate in atomic bonding. A simulation of this phenomenon requires an accurate description of atoms bonding at the crack tip, while at the same time including very large volumes of strained material. The resolution of information needed on a given scale varies with distance from the crack tip. Far away, it is adequate to use the equations of motion for a macroscopic averaged continuum field. With decreasing distance from the crack tip, singularities in the elastic field are cutoff by atomic scale phenomena and eventually by breaking electronic bonds. These phenomena on the one hand require more information to describe, but on the other hand, are dominant in

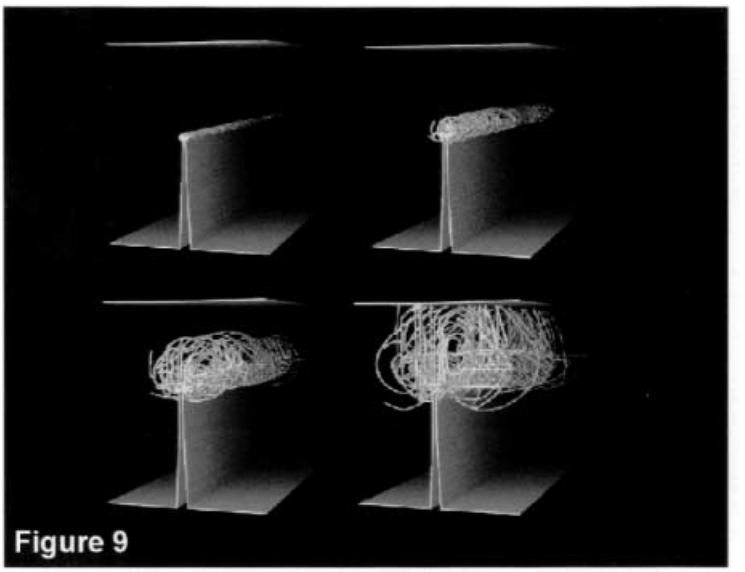

Fig. 9. Late-time sequence of the dislocations propagating through the 100 million atom slab shown for increasing times. With the exceptions that the entire depth and the top surface are shown, the rendering procedure described in Figure 8 applies.

successively smaller regions of materials. This suggests a natural physical domain decomposition in space: Angstrom cube volumes where electronic excitations must be considered explicitly, nanometer cube regions where atomic processes must be described, and micron cube and greater regions where displacement fields are sufficient. This spatial decomposition makes it possible to combine different simulation methods describing the different physical regions into a single powerful simulation tool.

We have recently presented a method which dynamically couples continuum mechanics far from the crack, empirical-potential molecular dynamics near the crack, and quantum tight-binding dynamics at the crack tip, to simulate fracture in silicon (Abraham et al., 1998). Continuum mechanics has long been fruitfully applied to the study of fracture, by explicitly putting in preexisting cracks or a phenomenological description of material decohesion. We use it to efficiently describe large parts of the system that are elastically deformed but do not include highly strained or broken bonds. Closer to the crack, as the strains become larger and the continuum description becomes less accurate, empirical potential molecular dynamics provides a fast atomistic description. While molecular dynamics can be used to simulate fracture by itself, the empirical potentials that govern the interaction between atoms may not reliably describe the breaking of bonds at the crack tip. A more accurate description in this region is given by tight-binding quantum mechanics, a method that simulates classical nuclei interacting via a simple quantum-mechanical description of the electrons that form

interatomic bonds. We describe the multiscale simulation method as MAAD. The multiscale MAAD approach should be considered only a beginning. Indeed, many improvements are currently being made (Abraham, et al., 2000).

Acknowledgments: I acknowledge support from the DOD Maui High Performance Computing Center, the USAF Wright-Patterson Computing Center and the San Diego Supercomputer Center and wish to thank DOD's HPCMO for a Grand Challenge award.

References

1. Abraham, F. F., Computational Statistical Mechanics: Methodology, Applications and Supercomputing, Advances in Physics, 35, 1 (1986). Concerning the supercomputing, this article is dated but still useful.
2. Abraham, F. F., Europhysics Letters, 38, No. 2, 103 (1997).
3. Abraham, F. F., Brodbeck D., Rafey R. and Rudge, W. E., Phys.Rev.Lett., 73, 272 (1994).
4. Abraham, F. F., Brodbeck, D., Rudge, W. E. and Xu, X-P, J.Mech.Phys.Solids, 45, 1595 (1997).
5. Abraham, F. F., Broughton, J. Q., Bernstein, N. and Kaxiras, E., Computers In Physics, 12, 538 (1998); Europhys.Lett., 44, 783-787 (1998).
6. Abraham, F. F., Bernstein, N., Broughton, J. Q. and Hess, D., MRS Bulletin, , (2000).
7. Abraham, F. F. and Gao, H., Phys.Rev.Lett., , (2000).
8. Abraham, F. F., Schneider, D., Land, B., Lifka, D., Skovira, J., Gerner, J. and Rosenkrantz, M., J.Mech.Phys.Solids, 45, 1461 (1997).
9. Aki, K. and Richards, P.G., Quantitative Seismology (W. H. Freeman, 2 Volumes, 1980).
10. Andrews, D.J., J. Geophys. Res., 81, 5679 (1976).
11. Archuleta, R. J., Bull. Seis. Soc., Am., 72, 1927 (1982).
12. Broberg, K. B., Int. J. Fract., 39, 1 (1989).
13. Broberg, K. B., Cracks and Fracture (Academic Press, San Diego, 1999).
14. Burridge, R., Geophys. J. Roy. Astron. Soc., 35, 439 (1973).
15. Clementi, E., Phil.Trans. Roy. Soc., A326, 445-470 (1988).
16. Freund, L.B., Dynamical Fracture Mechanics (Cambridge Univ. Press, New York, 1990).
17. Freund, L. B., J. of Geophys. Res., 84, 2199 (1979).
18. Gao, H., J. Mech. Phys. Solids, 41, 457 (1993).
19. Gao, H., Huang, Y. and Abraham, F. F., unpublished.
20. Fineberg, J., Gross, S., Marder, M. and Swinney H., Phys.Rev.Lett., 67, 457 (1991). Phys.Rev.B, 45, 5146 (1992).
21. Gordon, J. E., The New Science of Strong Materials (Princeton University Press, New Jersey, 1984).
22. Rosakis, A.J., Samudrala, O. and Coker, D., Science, 284, 1337 (1999).

Avalanches and Damage Clusters in Fracture Processes

Stefano Zapperi[1], Purusattam Ray[2], H. Eugene Stanley[3], and Alessandro Vespignani[4]

[1] INFM, sezione di Roma1, Dipartimento di Fisica, Universitá "La Sapienza", P.le A. Moro 2, 00185 Roma Italy
[2] The Institute of Mathematical Sciences, C. I. T. Campus, Chennai 600 113, India
[3] Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215
[4] The Abdus Salam International Centre for Theoretical Physics (ICTP) P.O. Box 586, 34100 Trieste, Italy

Abstract. By simulating two-dimensional models of electric breakdown and fracture it is possible to characterize the rupture of disordered materials subject to an increasing external stress. We provide a review of numerical and analytical results concerning the scaling properties of avalanche events prior the macroscopic breakdown of the material. The obtained results suggest a scenario that describes fractures as a first-order phase transition in the proximity of a spinodal-like instability. Finally, we discuss the properties of the avalanche branching ratio in the present context.

The understanding of the breakdown of disordered systems[1] has progressed to a large extent with new experiments on the precursors activity. The response of the material to an increasing external stress takes place in bursts distributed over a wide range of scales, and the main tool used in experimental studies is the measure of acoustic emission (AE) signal produced by microfractures occurring before the breakdown. Examples are found in the fracturing of wooden composite [2], cellular glass [3] and concrete [4], in hydrogen precipitation [5], in dislocation motion in ice crystals [6] and in volcanic activity [7]. The statistical analysis of these experimental observations points out the presence of scaling laws and poses many fundamental questions related to a description of fracture processes in terms of critical phenomena.

From a theoretical point of view, large scale simulations of lattice models are a very effective tools in the study of disordered materials breakdown. In these models a conductor is represented by a resistor network (scalar case) and the elastic medium by a spring network (vectorial case), where the disorder is modeled by random failure threshold or bond dilution. These models can be used in the quasistatic approximation in order to study the dynamical properties of systems that can be effectively considered at zero temperature and only the quenched disorder is relevant. This is the situation we have investigated in a series of papers[8,9]. In these works, we numerically studied the random fuse model [10,11,12,13] and a spring network [14,15]. Here we want to provide a review of these results indicating that the precursor behavior near the breakdown in

D. Reguera, L.L. Bonilla, and J.M. Rubí (Eds.): LNP 567, pp. 452–459, 2001.
© Springer-Verlag Berlin Heidelberg 2001

disordered systems is analogous to the formation of droplets observed close to a spinodal instability in first-order phase transitions[16].

For the sake of simplicity we shall consider the random fuse model, but the following results are valid also for the spring network. We simulate the random fuse model [10] on a tilted square lattice, with periodic boundary conditions in the transverse direction. To every bond i of the lattice we associate a fuse of unit conductivity $\sigma_i = 1$. An external current I or voltage V is then applied to the system by imposing suitable boundary conditions on two opposite edges of the lattice. When the current in the bond exceeds a randomly distributed threshold D_i the bond becomes an insulator ($\sigma_i = 0$). The voltage drops ΔV_i for each bond are computed solving the Kirchhoff equations for the network and the distribution of thresholds is chosen to be uniform in the interval $[1 - \Delta, 1 + \Delta]$. We impose an external current I through the lattice and we increase it at an infinitesimal rate. When a bond fails, we recompute the currents to see if other failures occur. The process is continued until a path of broken bonds spans the lattice and no current flows anymore.

We determine the cluster size distribution $n(s, I)$, which is defined as the number of clusters formed by s neighboring broken bonds when the applied current is I. The moments ($M_k(I) \equiv \int s^k n(s, I) ds$ is the k-th moment) of $n(s, I)$ describe much of the physics associated with the breakdown process. We determine $n(s, I)$ by averaging over the various threshold distribution configurations. The first moment $M_1(I)$ is the total number of broken bonds due to the current I and we measure the average $\langle m \rangle$ of the quantity $m = dM_1(I)/dI$ over various realizations of the process. In Fig. 1 we plot $\langle m \rangle^{-2}$ as a function of the reduced current I/I_c, where I_c is the average breakdown current, and we see that the graph is linear, signalling a divergence close to the breakdown as $(I_c - I)^{-1/2}$.

We also measure the distribution of avalanche sizes, integrating over all the values of the current. i.e. number of bonds m that break for a given value of the current, in a particular realization of the process.We have found that $P(m, I) \sim m^{-3/2} f(m(I_c - I))$, where I_c is the average breakdown current. By integrating over all the values of the current we get a *global* avalanche distribution $P(m) \sim m^{-5/2}$ (see Fig. 2). We have also checked that the cut-off of the distribution increases with the system size. A similar result for smaller lattice sizes ($L = 40$) was previously reported by Hansen and Hemmer [17], who also pointed out the similarity with the predictions of the Fiber Bundle Model (FBM).

A more detailed analysis can be performed by inspecting the higher moments of the cluster size distribution $n(s, I)$. The average cluster size $S \equiv M_2/M_1$ increases with I. However, by plotting S for different system sizes, we observe that the cluster size is not diverging (Fig. 3). To clarify this point, we confirm that $S(I_c)$ does not show scaling with the lattice size L. We also study the number of clusters $n_c \equiv M_0$ as a function of the current and for different system sizes We observe that n_c scales as

$$n_c = L^2 g(I/L), \tag{1}$$

which is expected for a first order transition. A first order behavior is also recovered by studying the behavior of the lattice conductivity in the model. We

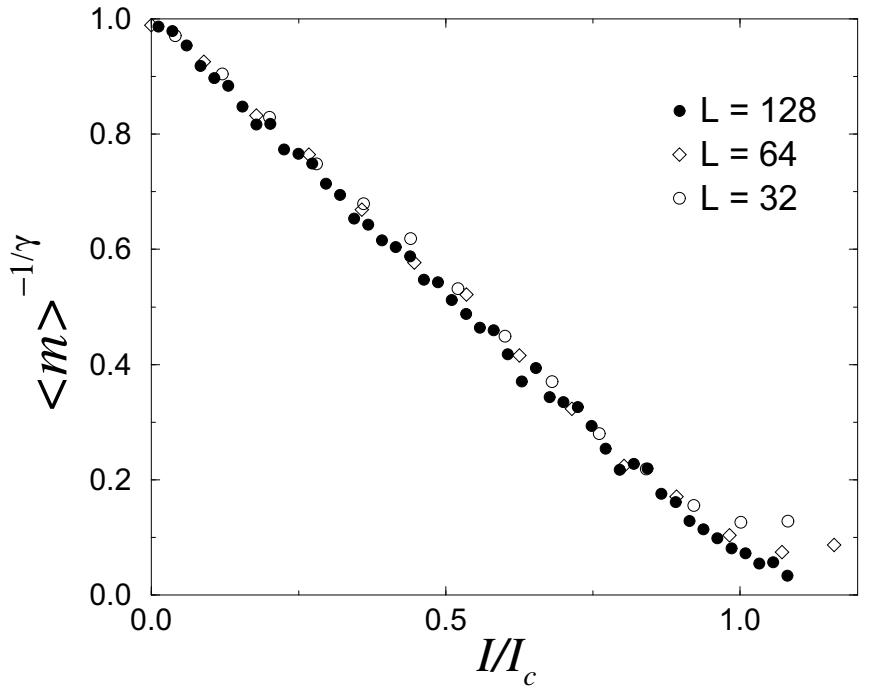

Fig. 1. The average avalanche size in the fuse model scaled with the mean-field exponent ($\gamma = 1/2$) as a function of I/I_c, for different values of the system size L. The linearity of the plot supports the validity of the mean-field calculations.

plot the conductivity averaged over different realizations of the disorder and we observe a smooth curve with a slope at the breakdown that becomes sharper as the system size increases (Fig. 4).

In order to gain some analytical insight on the previous results, we can generalize the analysis of Ref. [20] to derive a simple mean-field theory for fracture in disordered media. To illustrate the mean-field theory we will consider for simplicity the random fuse model as previously described, by recasting the dynamics of the model in terms of the externally applied current I. We can write the energy of the lattice, in full generality, as

$$E(I, \{\sigma\}) = \frac{1}{2}\left(\frac{I^2}{G(\{\sigma\})} - \sum_i \sigma_i D_i^2\right), \tag{2}$$

where $G(\{\sigma\})$ is the total conductivity of the lattice and is a complicated function of the local conductivities. We can estimate $G(\{\sigma\})$ using the effective medium theory [19], which in our case gives

$$G(\{\sigma\}) = 2\phi - 1, \tag{3}$$

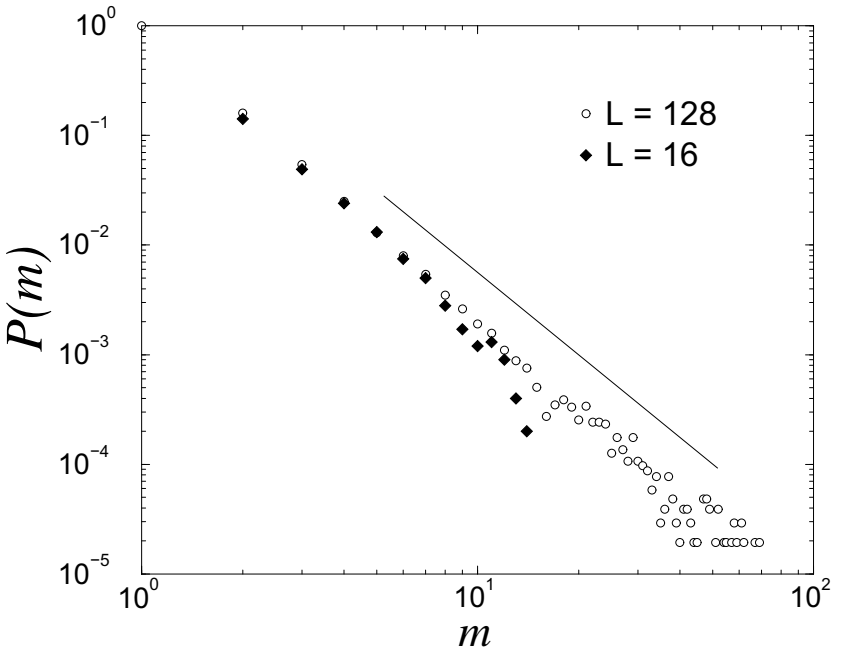

Fig. 2. The avalanche size distributions in the fuse model for two values of system size plotted in log-log scale. A line with the mean-field value $\tau' = 5/2$ of the exponent is plotted for reference.

where $\phi \equiv \sum_i \sigma_i / L^2$. The total energy can then be written as a sum over the local bonds driven with an effective random fields h_i

$$E_{MF}(I, \{\sigma\}) = \sum_i \sigma_i h_i = \frac{1}{2} \sum_i \sigma_i \left(\frac{I^2}{L^2 \phi (2\phi - 1)} - D_i^2 \right). \tag{4}$$

As shown in Ref.[8], the value of ϕ can be computed self-consistently obtaining a solution only if $I < I_c$. Close to I_c, ϕ scales as $\phi - \phi_c \sim (I_c - I)^{1/2}$, and the average avalanche size $\langle m \rangle$ that is proportional to the "susceptibility" $d\phi/df$ [21] diverges at the breakdown as

$$\langle m \rangle \sim (I_c - I)^{-\gamma} \qquad \gamma = 1/2. \tag{5}$$

We can also obtain the mean-field avalanche size distribution from the exact results derived for the FBM [18]

$$P(m, I) \sim m^{-\tau} f(m(I_c - I)^{\kappa}); \qquad \tau = \frac{3}{2}, \quad \kappa = 1, \tag{6}$$

where m is the number of bonds that break as function of the current. Eq. (6) can also be obtained in the case of bond damage using similar arguments. The

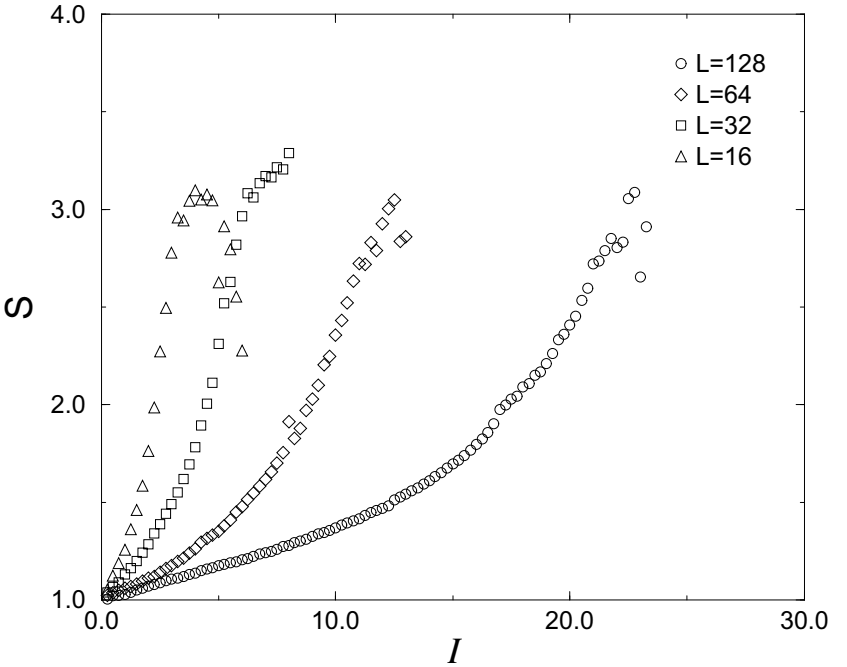

Fig. 3. The average cluster size as a function of the current for different system sizes. Note that the cluster size does not diverge.

analytical predictions obtained are in perfect agreement with the numerical findings we have shown on the fuse network and are confirmed in a spring network model that takes into account the vectorial nature of fracture processes[8].

The above results clarify the nature of the breakdown process in the presence of quenched disorder. We have found that the breakdown is preceded by avalanches distributed as power laws, and that the scaling exponents are in quantitative agreement with the prediction of mean-field calculations. However, we have also found that only *globally* defined quantity such as $\langle m(I) \rangle$ and $P(m)$ display scaling, while locally defined quantities, such as S, do not show any singular behavior as we would expect for second order critical behavior. These results can be reconciled with the scaling behavior of the precursor avalanche distribution by invoking the occurrence of a first order transition close to a spinodal point[16]. The elastic state is considered to be metastable, as soon as a non-zero stress is applied. Due to the presence of disorder, the system evolves through a series of metastable states towards the final instability. This occurs with the nucleation of cracks growing up to a critical size s_c at which they *coalesce* forming the macroscopic crack. Contrary to percolation, in this case there is no incipient spanning cluster prior to rupture. What explains then the scaling in the avalanche statistics and the susceptibility? We recall that elastic (or electric) forces are long-range. When nucleation occurs close to a spinodal instability—

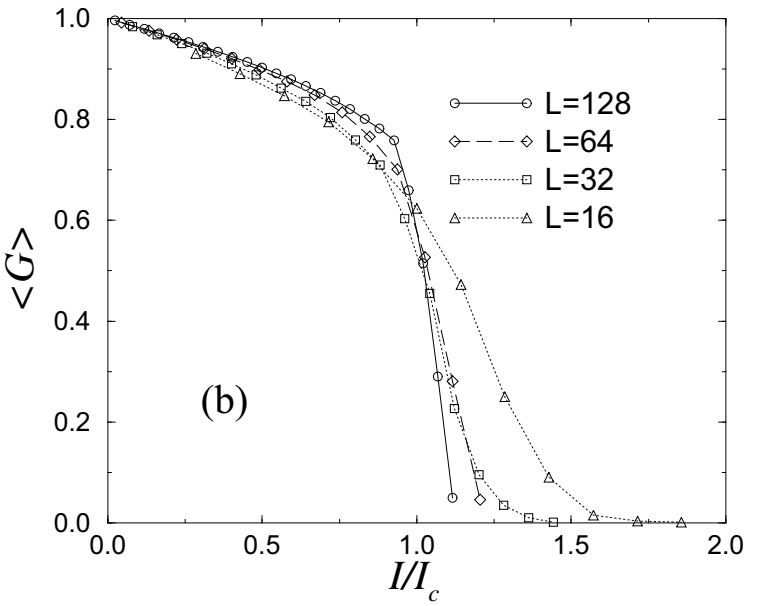

Fig. 4. The conductivity as function of I/I_c averaged over different realizations of the disorder. Note that the discrete jump, indicative of a first-order transition, is smoothed for small system sizes.

which is well defined only for mean-field or long-range interactions—one expects a divergent susceptibility [16]. This is not naively related to the fluctuation of a geometrical quantity such as the crack size, which is not diverging at the spinodal. In order to describe geometrically the susceptibility, it is necessary to define the clusters in a peculiar way, considering each site connected with all the others within the range of interactions [22]. These fluctuations are therefore different from those encountered in a second-order phase transition.

The interpretation we gave has been recently questioned studying the behavior of the avalanche 'branching ratio' [23,24]. It has been proposed to identify the branching ratio with the order parameter of the transition. The fact that $R \to 1$ at breakdown, was then used to claim that the transition was of second order type. We notice first that the analysis was based on an incorrect definition of a branching ratio. In a branching process [25], at each time step an active site can give rise to n offsprings with probability p_n, where $\sum_n p_n = 1$, or die with probability p_0. The branching ratio is then defined as $R = \sum_n np_n$, so that for $R < 1$ the process stops with probability one in the limit $t \to \infty$ while for $R > 1$ it runs forever. The 'critical' point is associated with $R = R_c = 1$. In the *subcritical* case (i.e. $R < 1$) one can easily compute the average 'avalanche' size $\langle m \rangle$, defined as the total number of active sites in the process, as $\langle m \rangle = 1/(1-R)$. In Refs. [23,24] this result was instead used as a definition, inverting the relation:

$R = 1 - 1/\langle m \rangle$. Since this relation is valid only for $R < 1$, it can not be used to study the transition (i.e. when R is supposed to cross the 'critical point' $R = 1$!).

In addition, analyzing correctly the behavior of R can not give further information on the nature of the transition with respect to the analysis discussed here. From Eq. (5) one can immediatly conclude that $R \sim 1 - (I_c - I)^{1/2}$, so that $R \to 1$ for $I \to I_c$. This is just another way to state that the avalanche size is diverging at the breakdown point and it is not in contrast to the spinodal first-order transition we are advocating. One should always keep in mind that the spinodal point is a kind of critical point, thus one can see, for instance, the divergence of the quasi-static susceptibility χ [22]. Following Ref. [24], we could define the branching ratio in the long-range Ising model [22] as $R \equiv 1 - 1/\chi$ and find that when the field H reaches the spinodal value H_c, $R \to 1$. However, the phase transition is still first-order, since the magnetization changes abrubtly sign. The answer to these apparent paradoxes is simple: the branching ratio *is not the order parameter of the transition.*

It is, however, worth discussing some "*caveat*" to the present discussion since in driven disordered systems the notions of metastability, spinodal point and nucleation are not well defined. In particular, the identification of ϕ and I with the order and control parameters is justified only in MF theory. In two dimensions, simple homogeneous scaling fails in the presence of disorder. The breakdown current I_c has a logarithmic size dependence [11],

$$I_c \sim \frac{L}{\log L}, \tag{7}$$

which cannot be interpreted in MF theory. A similar dependence is also present in the fraction of broken bond before breakdown ϕ_c. While for finite systems the picture we have presented is completely consistent, it is not obvious how to perform the $L \to \infty$ limit. In order to obtain intensive parameters in this limit one should rescale I and ϕ by an appropriate logarithmic factor. This can be done implicitly analyzing the data in terms of I/I_c, as in Ref. [2]. Notice that this problem occurs even if one would like to interpret breakdown as a genuine second-order transition- Finally, we note that in most cases the final breakdown starts from existing defects in the material that we have not contemplated in our system.

References

1. K. K. Bardhan, B. K. Chakrabarti and A. Hansen (eds.), *Non-linearity and breakdown in soft condensed matter*, (Springer Verlag, Berlin, 1994). B. K. Chakrabarti and L. G. Benguigui, *Statistical physics of fracture and breakdown in disordered systems* (Oxford Univ. Press, Oxford, 1997).
2. A. Garcimartín, A. Guarino, L. Bellon and S. Ciliberto, Phys. Rev. Lett. **79**, 3202 (1997); A. Guarino, A. Garcimartín and S. Ciliberto, Eur. Phys. J. B **6**, 13 (1998).
3. C. Maes, A. Van Moffaert, H. Frederix and H. Strauven, Phys. Rev. B **57**, 4987 (1998).

4. A. Petri, G. Paparo, A. Vespignani, A. Alippi and M. Costantini, Phys. Rev. Lett. **73**, 3423 (1994).
5. G. Cannelli, R. Cantelli and F. Cordero, Phys. Rev. Lett. **70**, 3923 (1993).
6. J. Weiss and. J.-R. Grasso, J. Phys. Chem. B **101**, 6113 (1997).
7. P. Diodati, F. Marchesoni and S. Piazza, Phys. Rev. Lett. **67**, 2239 (1991).
8. S. Zapperi, P. Ray, H. E. Stanley and A. Vespignani, Phys. Rev. Lett. **78**, 1408 (1997); Phys. Rev. E **59**, 5049 (1999)
9. S. Zapperi, A. Vespignani and H. E. Stanley, Nature **388**, 658 (1997).
10. L. de Arcangelis, S. Redner and H. J. Herrmann, J. Phys. Lett. (Paris) **46**, L585 (1985).
11. P. Duxbury, P. D. Beale and P. L. Leath, Phys. Rev. Lett. **57**, 1052 (1986).
12. L. de Arcangelis and H.J. Herrmann, Phys. Rev. B **39**, 2678 (1989).
13. B. Kahng, G. G. Batrouni, S. Redner, L. de Arcangelis and H. J. Herrmann, Phys. Rev. B **37**, 7625 (1988).
14. P. Ray and G. Date, Physica A **229**, 26 (1996).
15. K.-T. Leung and J. V. Andersen, Europhys. Lett. **38** 589 (1997); K.-T. Leung, J. V. Andersen and D. Sornette, Phys. Rev. Lett. **80**, 1916 (1998).
16. C. Unger and W. Klein, Phys. Rev. B **29**, 2698 (1984); *ibidem* **31**, 6127 (1985). For a review see L. Monette, Int. J. of Mod. Phys B **8**, 1417 (1994).
17. A. Hansen and P. C. Hemmer, Phys. Lett. A **184**, 394 (1994).
18. P. C. Hemmer and A. Hansen, J. Appl. Mech. **59**, 909 (1992); M. Kloster, A. Hansen and P. C. Hemmer Phys. Rev E **56**, 2615 (1997).
19. S. Kirkpatrik, Rev. Mod. Phys. **45**, 574 (1973).
20. R. L. B. Selinger, Z.-G. Wang, W. M. Gelbart and A. Ben-Saul, Phys. Rev. A **43**, 4396 (1991); Z.-G. Wang, U. Landman, R. L. B. Selinger and W. M. Gelbart, Phys. Rev. B **44** 378 (1991).
21. M. Acharaya and B. K. Chakrabarti, Phys. Rev. E **53**, 140 (1996); M. Acharaya, P. Ray and B. K. Chakrabarti, Physica A **224**, 287 (1996).
22. D. Heerman, W. Klein and D. Stauffer, Phys. Rev. Lett. **49** 1262 (1982); T. Ray and W. Klein, J. Stat. Phys. **61**, 891 (1990).
23. G. Caldarelli, C. Castellano and A. Petri, Physica A **270**, 15 (1999).
24. Y. Moreno, J. B. Gomez, A. F. Pacheco, Phys. Rev. Lett. **85**, 2865 (2000)
25. T. E. Harris, *The Theory of Branching Processes* (Dover, New York, 1989).

Index

Avalanches (45.70.Ht): IV.6
Boundary layer and shear turbulence (47.27.Nz): I.2, I.3, I.6
Chaos (47.52.+j): I.8
Chemically reactive flows (47.70.Fw): II.4
Compressible flows; shock and detonation phenomena (47.40.-x): II.3
Control of chaos, applications of chaos (05.45.Gg): I.8
Dynamics (capillary waves) (68.03.Kn): III.9
Ecology and evolution (87.23.-n): III.3
Electric discharges (52.80.-s): III.5
Fatigue, brittleness, fracture, and cracks (62.20.Mk): IV.5
Fatigue, corrosion fatigue, embrittlement, cracking, fracture and failure (81.40.Np): IV.5
Fracture mechanics, fatigue and cracks (46.50.+a): IV.5, IV.6
Granular flow: mixing, segregation and stratification (45.70.Mg): IV.1, IV.2, IV.3, IV.4
Granular systems (45.70.-n): IV.1, IV.2, IV.3, IV.4
Growth from melts; zone melting and refining (81.10.Fq): III.6
Growth in microgravity environments (81.10.Mx): III.6
Hydrodynamic stability (47.20.-k): II.2
Hydrodynamic waves (47.35.+i): III.9
Lagrangian and Hamiltonian mechanics (45.20.Jj): I.8
Liquid crystals (42.70.Df): III.7
Magnetized plasmas (52.25.Xz): III.8
Magnetohydrodynamics and electrohydrodynamics (47.65.+a): III.10
Nanoscale pattern formation (81.16.Rf): III.4
Nonlinear dynamics and nonlinear dynamical systems (05.45.-a): II.1
Pattern formation (45.70.Qj): IV.1
Pattern formation in reactions with diffusion, flow and heat transfer (82.40.Ck): III.4
Patterns (89.75.Kd): III.5
Pattern selection; pattern formation (47.54.+r): I.4
Pattern selection; pattern formation (47.54.+r): III.7
Plasma turbulence (52.35.Ra): III.8
Population dynamics and ecological pattern formation (87.23.Cc): III.1, III.2, III.3
Reactions in flames, combustion, and explosions (82.33.Vx): II.1, II.2, II.3
Spatiotemporal pattern formation in cellular populations (87.18.Hf): III.1, III.2
Static sandpiles; granular compaction (45.70.Cc): IV.1, IV.4
Temporal and spatial patterns in surface reactions (82.40.Np): III.4
Theory and modeling; computer simulation (87.15.Aa): III.1
Transition to turbulence (47.27.Cn): I.4, I.7
Turbulence simulation and modeling (47.27.Eq): I.1, I.3, I.7
Turbulent flows, convection, and heat transfer (47.27.-i): I.5, I.6
Vortices and turbulence (67.40.Vs): I.1, I.2, I.3, I.5
Wave propagation, fracture, and crack healing (83.60.Uv): IV.5

List of Participants

F. Abraham
IBM Almaden Research Center, USA
T. Alarcón
Universitat de Barcelona, Spain
S. Alonso
Universitat de Barcelona, Spain
E. Alvarez
Universitat de Barcelona, Spain
A. Arenas
Universitat Rovira i Virgili, Spain
M. Arrayas
U. Leiden, The Netherlands
R.P. Behringer
Duke University, USA
E. Ben-Jacob
Tel-Aviv University, Israel
K. Berkouk
University of Exeter, United Kingdom
A. Bohn
Darmstadt Univ. Technology, Germany
L.L Bonilla
U. Carlos III, Spain
T. Borzsonyi
Universites Paris VII et Paris VI, France
J. Brey
U. Sevilla, Spain
A. Buka
Research Institute for Solid State Physics and Optics, Hungary
O. Bulashenko
Universitat de Barcelona, Spain
J. Carlson
University of California at Santa Barbara, USA
A. Carpio
Universidad Complutense de Madrid, Spain

M. Carretero
U. Carlos III, Spain
J. Cartwright
CSIC-Universidad de Granada, Spain
J. Casademunt
Universitat de Barcelona, Spain
M. Castellà
Secretary
S. Chountasis
I. for Scientific Interchange, Italy
P. Clavin
Université d'Aix Marseille, France
A. Corral
Universitat Politècnica de Catalunya, Spain
Y. D'Angelo
LCD ENSMA-CNRS, France
E. Degroote
U. Politécnica Madrid, Spain
C. De Miguel
Universitat de Barcelona, Spain
A. Demircan
University of Potsdam, Germany
C. Dopazo
U. de Zaragoza, Spain
U. Ebert
Leiden University, The Netherlands
S. Esipov
Jersey City, USA
D. Ewing
McMaster University, Canada
J. Fajans
U.C. Berkeley, USA
H. Fenichel
U. of Cincinnati, USA
J. Fernández
Universitat de Barcelona, Spain
J.A. Ferré Vidal
U. Rovira i Virgili, Spain
R. Folch
Universitat de Barcelona, Spain
P.L. García Ybarra
CIEMAT, Spain
L. Garrido
Universitat de Barcelona, Spain

A.A. Garriga
Universitat de Barcelona, Spain
M.E. Glicksman
Rensselaer Polytech Institute, USA
D. Gomila
I. Mediterrani Estudis Avançats, Spain
M. González
Secretary
F. Guinea
ICMM, CSIC, Spain
A. Hernández
Universitat de Barcelona, Spain
F.J. Higuera
U. Politécnica Madrid, Spain
J. Jiménez
Universidad Politécnica de Madrid, Spain
G. Joulin
CNRS, France
V. L. Karbivs'kii
Academy of Sciences, Ukraina
S. Kida
National Institute for Fusion Science, Japan
M. Kindelan
U. Carlos III, Spain
A.M Lacasta
Universitat Politècnica de Catalunya, Spain
J.S. Lamb
Imperial College, London, United Kingdom
A. Liñán
Universidad Politécnica de Madrid, Spain
J.J. López Velázquez
U. Complutense (Madrid), Spain
P. Maini
University of Oxford, United Kingdom
P. Manneville
CNRS-UMR, France
M. Mayorga
U. A. Estado de México, Mexico
J.D. Mellado
Carlos III, Spain
A. Mikhailov
Fritz-Haber-Institut der Max-Planck-Gesellschaft, Germany
J. Mueller
Leiden University, The Netherlands

A. Muriel
U. of Philippines, Philippines
G. Oron
Universitat de Barcelona, Spain
I. Pagonabarraga
Universitat de Barcelona, Spain
R. Pastor
Universitat de Barcelona, Spain
E. Pauné
Universitat de Barcelona, Spain
I. Pérez
Universitat de Barcelona, Spain
A. Pérez-Madrid
Universitat de Barcelona, Spain
O. Piro
U. Illes Balears, Spain
G. Platero
ICMM, CSIC, Spain
T Poschel
Humboldt-Universitaet zu Berlin, Germany
I. Ramírez de la Piscina
Universitat Politècnica de Catalunya, Spain
J.M. Redondo
Universitat Politècnica de Catalunya, Spain
I. Rodríguez Cantalapiedra
Universitat Politècnica de Catalunya, Spain
M.A. R. de Cara
CSIC, Madrid, Spain
D. R. Pérez
UNED, Spain
A. Radosz
University of Technology, Poland
M. Ragwitz
Max Planck Institute, Germany
D. Reguera
Universitat de Barcelona, Spain
F. Ritort
Universitat de Barcelona, Spain
A. Rivero
ETSIA/ U. Politécnia Madrid, Spain
L. Romero-Salazar
U. A. Estado de México, México
A. Rosowsky
CEA, France

J.M. Rubí
Universitat de Barcelona, Spain
S. Ruediger
University of Potsdam, Germany
M.J. Ruiz-Montero
U. Sevilla, Spain
F. Sagués
Universitat de Barcelona, Spain
M. Sales
Universitat de Barcelona, Spain
C. Salueña
Universitat de Barcelona, Spain
O. Sánchez
U. de Granada, Spain
I. Santamaría
Universitat de Barcelona, Spain
T.E. Schilhabel
University of Southampton, United Kingdom
G. Searby
Université d'Aix Marseille, France
S. Skurativsky
Subbotin Inst. of Geophisics, Ukraiane
A.J. Smits
Princeton University, USA
C. Storm
Leiden University, The Netherlands
C. Treviño
UNAM, México
J.M. Truffaut
IRPHE (Marseille), France
J.M. Vega
U. Politécnica Madrid, Spain
A. Vespignani
ICTP, Italy
J.M. Vilar
Princeton, USA
F. Zoueshtiagh
School of Engineering, Coventry, United Kingdom